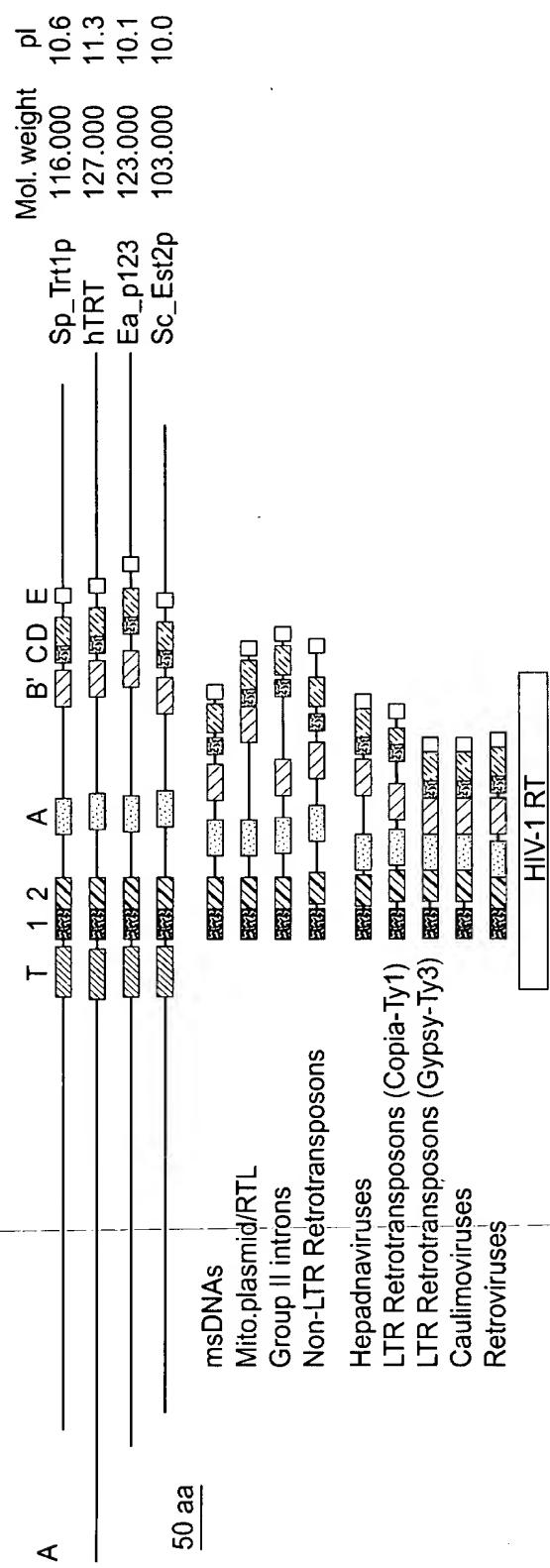


	Motif 0				Motif 1				Motif 2				Motif 3 (A)			
human		AKFLHWMISVYVVELLRSFFYVTTETTFQKNR			LFFYRKSVWSKLQSIGIROQLKRVQLRDVSEAEVRQHREARPALLTSRLRFIPKP--DGL				RPIVNMDYVVGARTFRREKRAERLTSRVKALF-SVLYNERYA				KDDLKHRMFGR-KKYFVRIDIKSCYDRIKQDLMFRIVKK-KLDPEFVIRKYATIATS			
tez1		ISETIEWLVLGKRSNAKMCLSDBEKRQKQIFAEFIYWLNSFLIPLQSFYITTESSDLRNR			TVYFRKDIIWKLLCRPFIT-SMMMEAFEKINENNVRMDTQK-TTLPPAVIRILLPKK--NTF				RLLTN-LRKRFLLIKMGNSNKKMIVSTNOTLRPVASILKHLINNEESSGIPENILEVYMKLLTF				KQRLLKKFNNVLPLEYFMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFVRSQYFFNTN			
EST2		LKDFRWLFISD---IWFTKHNPENLNQLAICFISWLFRQLIPKLIQTFYCTEISSTVT-			IVYFRHDITWNKLITPPFIVEYFKTYLVENNCRNHNSYTLS--NFNHSKMRILIPKKSNNEF				RIIAIPCRGADEEEFTIYKENHNKAIQPTQKILEYLRNKRPTSTFTKISPTQIADRRIKEF				EEFVCKWKQVGQPQLFFATMDIEKCYDSVNREREKLSTFLKTTKLSSDFWIMTAQILRKRN			
p123		TREISWMQVET-SAKHFYYFDHEN-IYVLWKLRLWFEDLVSLIRCFYYVTEQQKSYSK	*	*	*	*	*	*	RPIIMTFNKKIVNSDRKTTKLTTNTKLLNSHMLXTLKN-RMFKDPPGFAVFTNYDDVMKRY	*	*	*	*	*	*	*
		*	*	*	*	*	*	*								

FIG. 1

+



+

+

3/103

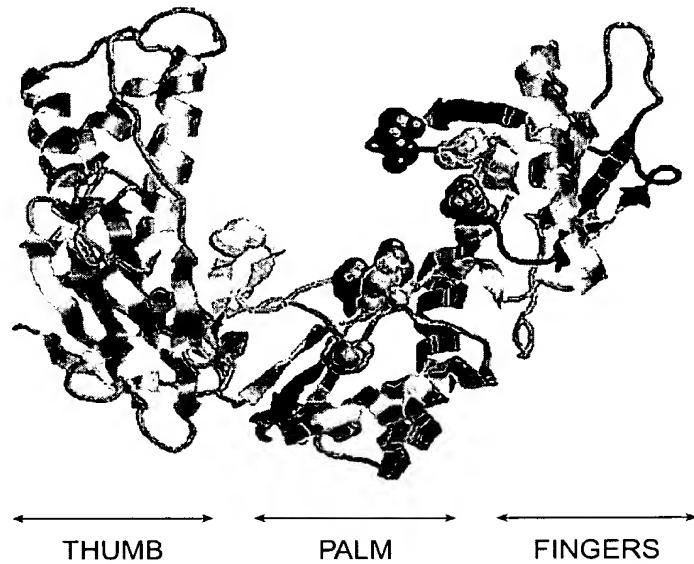


FIG. 3

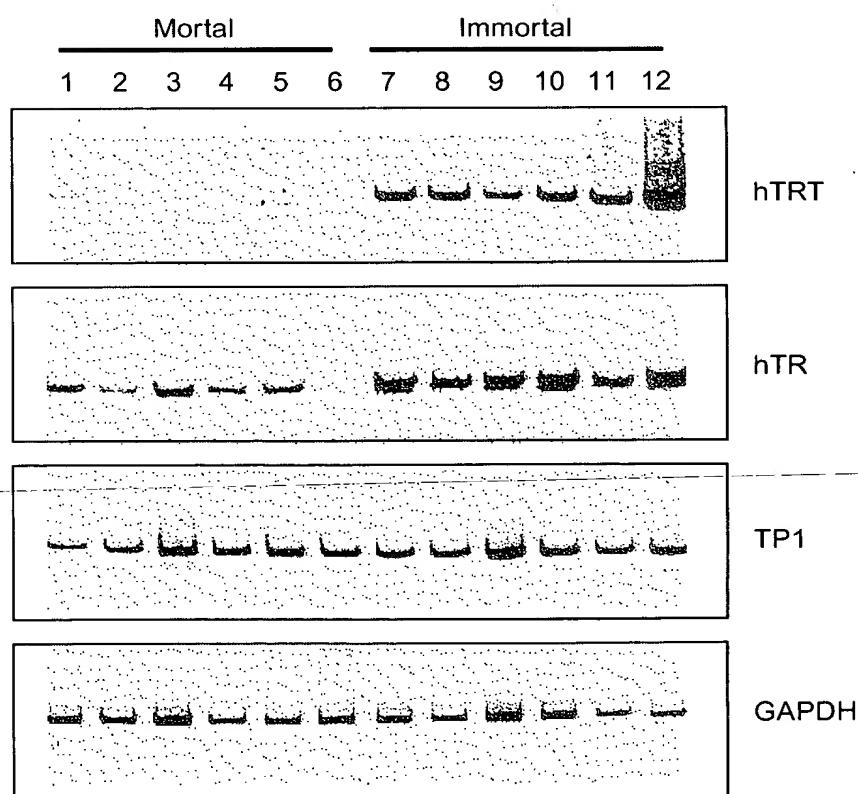


FIG. 5

+

		Motif T							
TRT	con	WL	hh	pffy	te	p	y	RK	W
<u>SP</u>	<u>Trt1p</u>	429	WLYNSFLIPILOQFFYITESSSLRNRRTVYFRKDIIWLLCRPFTSMKM						8
<u>bTRT</u>		546	WLMSVTVELLRSFFYVTTFQKNRLFFYRKSVWSKLQSIGIRQLK						10
Ea	p123	441	WIFEDLIVSLSIRCFFYVTEQQSYSKTTYYRKNIWWDVIMKMSIAIDLKK						8
<u>SC</u>	<u>Est2p</u>	366	WLFRQLIPKLIQTFFYCTEISSVT-IIVYFRHDTMNKLITPFIVEYFK						8

	Motif 1	Motif 2	Motif A
TRT con	h	FRHI h	PcLYFh
SP_trt1p	hRHIPKK p	FRITNLRKRFLLIKMGNSKMKLVNSTNQTL	FGRKKYFVRIDLTKSCYDRIKQDLMFRIVKKLKD
bTrt	NNVRMDTQKTTLPPAVIRLLPKNT-	0	82
EVROHREARPAILTSLRRFIPKPDG-	0	PPPELYFVKVDTVGAYDTIPODRLEVIASI1KPK	87
KEVEEWKSSLGFAPGKLRLIPKKTT-	0	GOPKLFATMDIEKCYDSVNREKLSTFLKTKL	100
CRHNHSYTLSNFHSSKMRIIPKSSNN	1	1	68
		VPPELYFMKF DVKSCYDSDIPRMCECMRILKDALKN	
		AF	
	R		
	hR h	h	h
	p hh h K	hDh	GY h
	LSNEIIGTGKEKKPMPRTVNIKPKGGS	27	EGGSNNWTEEVDLKKCFDTISHDL LIKELKRYISD
a1	0	0	20
Sdm_TART	SILRIGYYPPDAWKHAQVMKMLPKGKS	32	RKEYCSAFLFDLSEAFDRVHEGGLLKLAKILP
HIV-1	6	0	25
	EGKISKIGGPENPVNTPVFAIKKDST	1	LKKKSSTVVLGDAYFSVPLDEDFRKYTAFTIP
		1	7

	Motif B'	Motif C	Motif D	Motif E
TRT con	K Y Q GIPOGS LS hL h Y DL F	LLRL DDFLHIT	A F h G c p N CK	W G S
SP_Trt1p	SOYIQLQKVGIQPGSILSSFLCHFYMEDLIDDEYLSTFT	6 LLRVVDDFLFITVNKKD 0	AKKFLNLNLSLRGFEKHNFNSTSLEXTVI	17 KKMRPFFGFSV 181
hTrR	KSYVQCQGIPQGSILSLTLLCSCLYGDMENKLFGAI	5 LLRLVDDFLLLVTPHLTH 0	AKTFLRLTVRGVEYYGCVVNLRTVV	19 HGLFPWCGLL 197
Ea_p123	KFYKQTKGIGIPQGLCVSSILSFSFYATLEESSLGFL	14 LMRLTDDYLITLITQENN 0	AVLFIEKLLINVSRENGFKFMKQLQT	23 QDYCDWIGTSI 179
Sc_Est2p	KCYIREDGILFGQGSSLSSAPIVDLVYDDLIEFYSEFK F	8 ILKLAADDFLIISTDQQQ 0	VINIKKLMGGFQKYNAKANRDKILA	20 KELEVWKHSST 146
RRT con	hPQG pp hh h	h Y DDhhh	Gh h ck h	hlG h
Sc_a1	TYHKPMGLPQGSLISPILCNIVMLIVDNWLEDYI	55 YVRYADDILIGVIGSKN 2	KMICKDLNNFLNS - LGLTWNNEEKTLLI	4 ETPARFLGYN
Dm_TART	RAGOIIAGVGPQGSNLGPILYSIFSSDMELPHIVHP	7 LSTYADDITVLSIDLIA 6	NENYLKTFSDWADKGWISVNAAKTGH	25 ESKOSYLVGIL
HIV-1	GIRYQINVLPQWGKGSPLAFQSSMTKILEPFFKKQN	4 IYQYMDDLVYGSDELIG 1	HRTKIEELRQHLLRWGLTTPDKKHQK	0 EPPFLWNGITL

FIG. 4

+

5/103

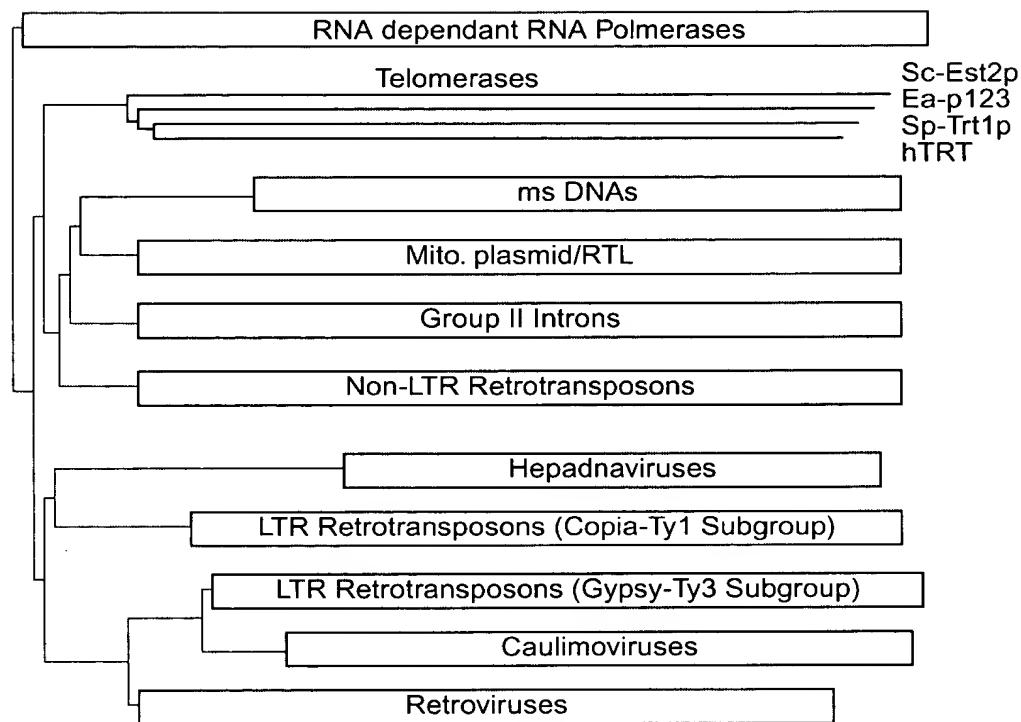


FIG. 6

+

+

6/103

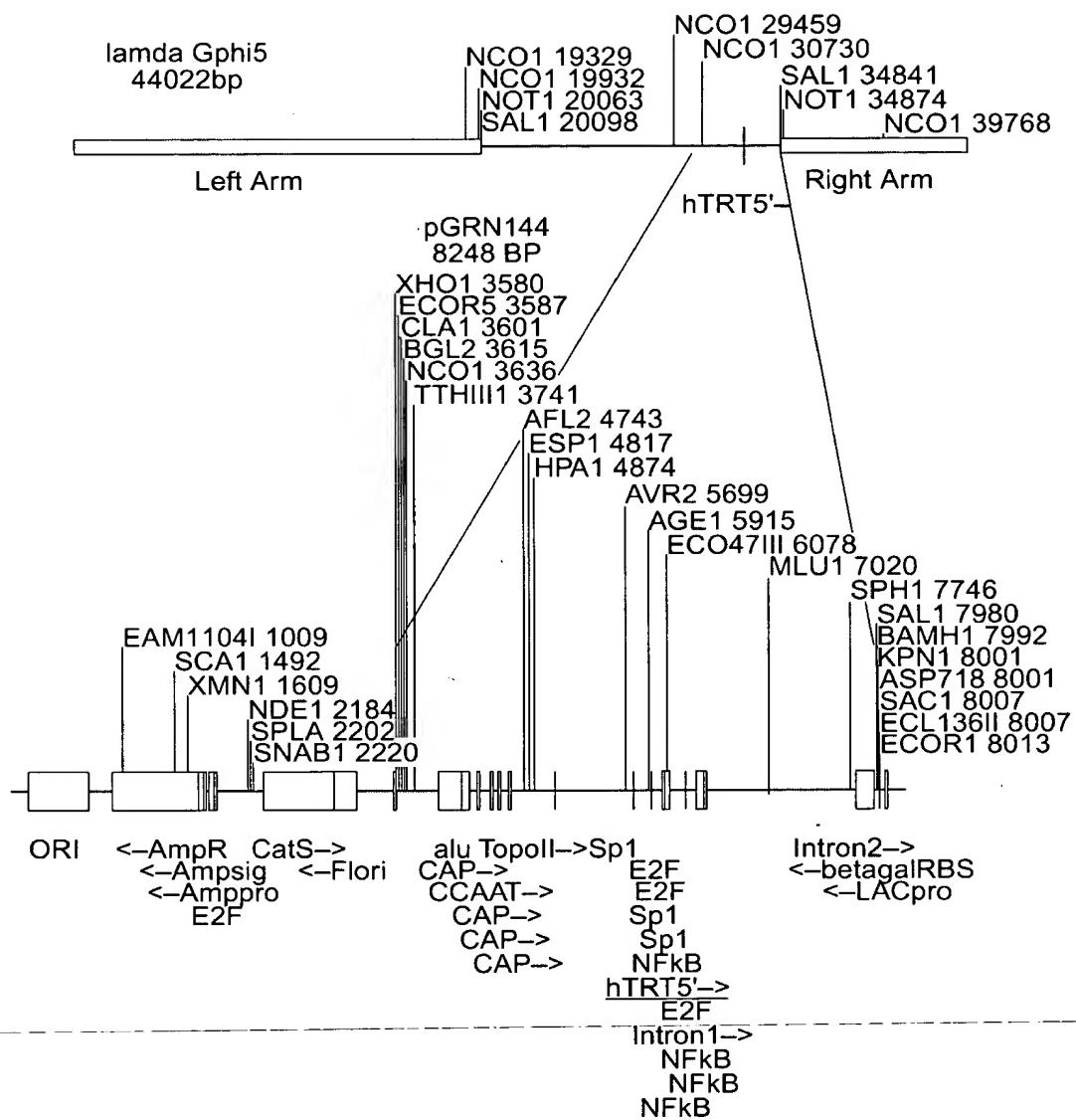


FIG. 7

+

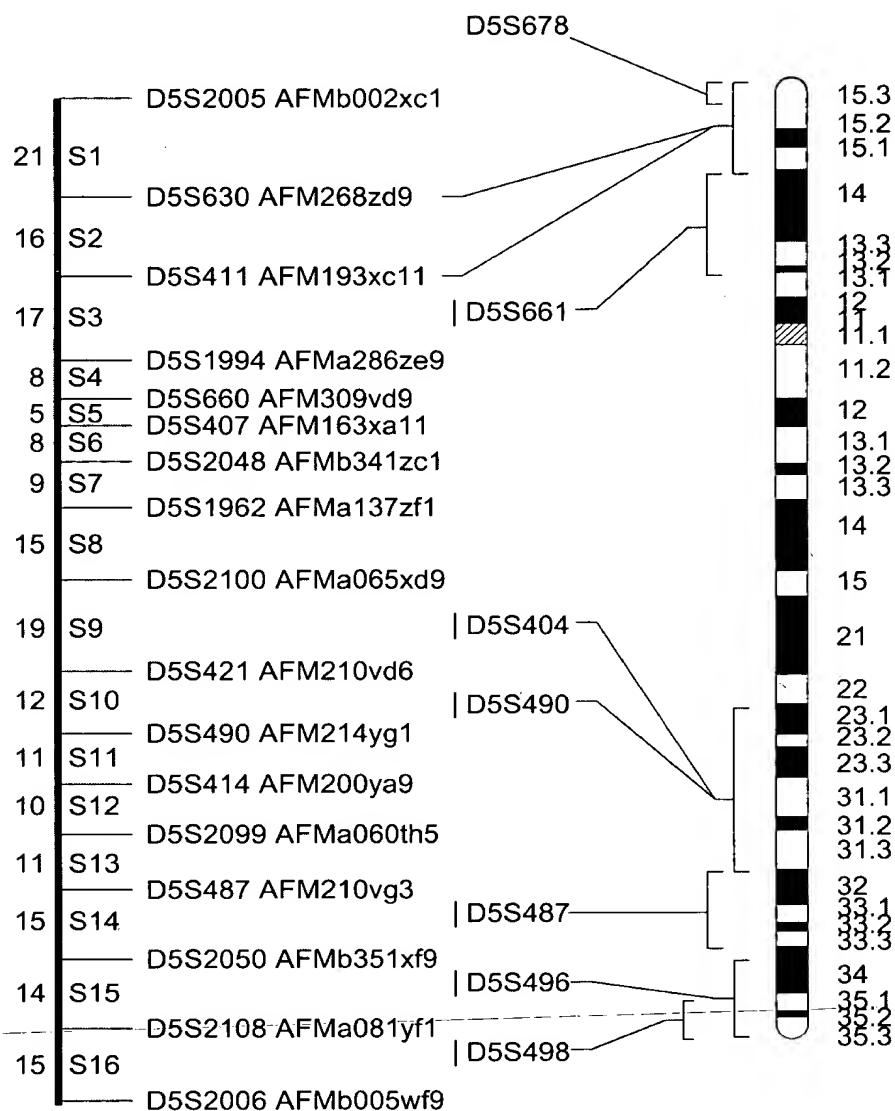


FIG. 8

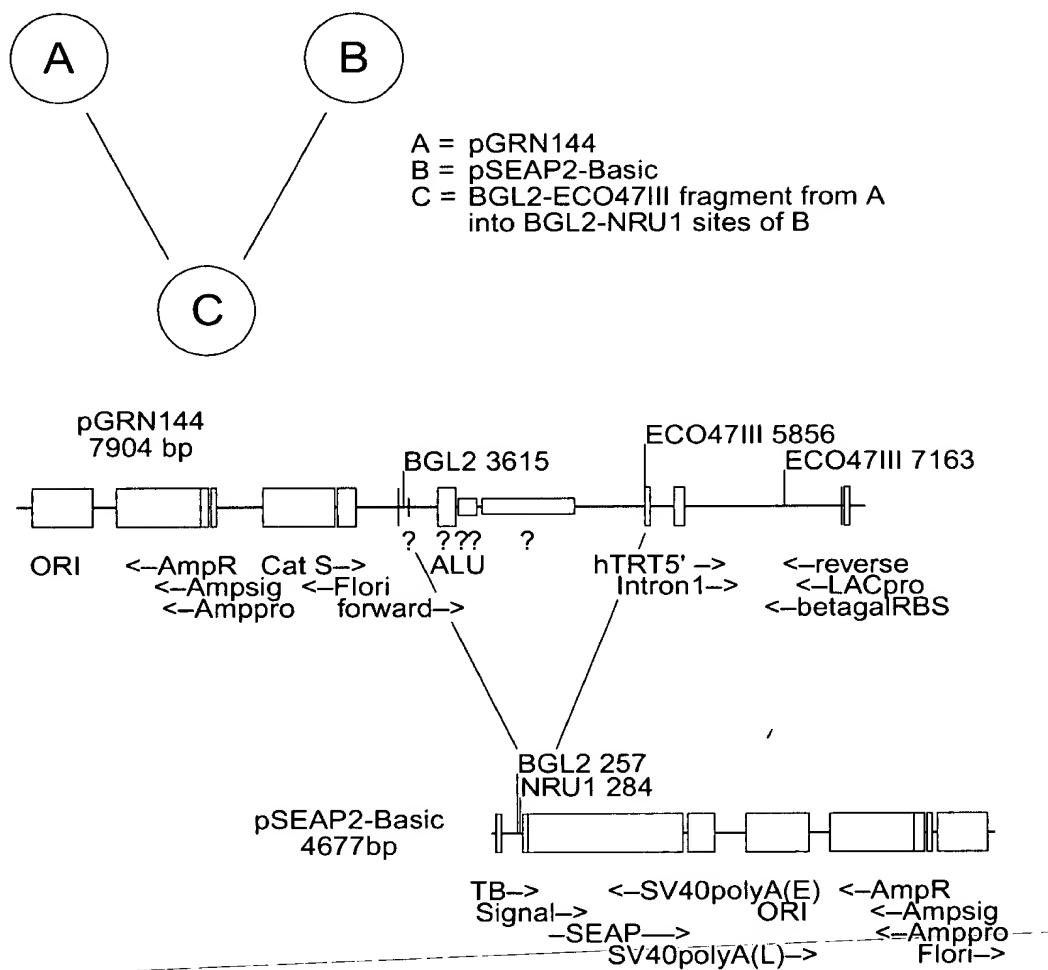


FIG. 9

+

9/103

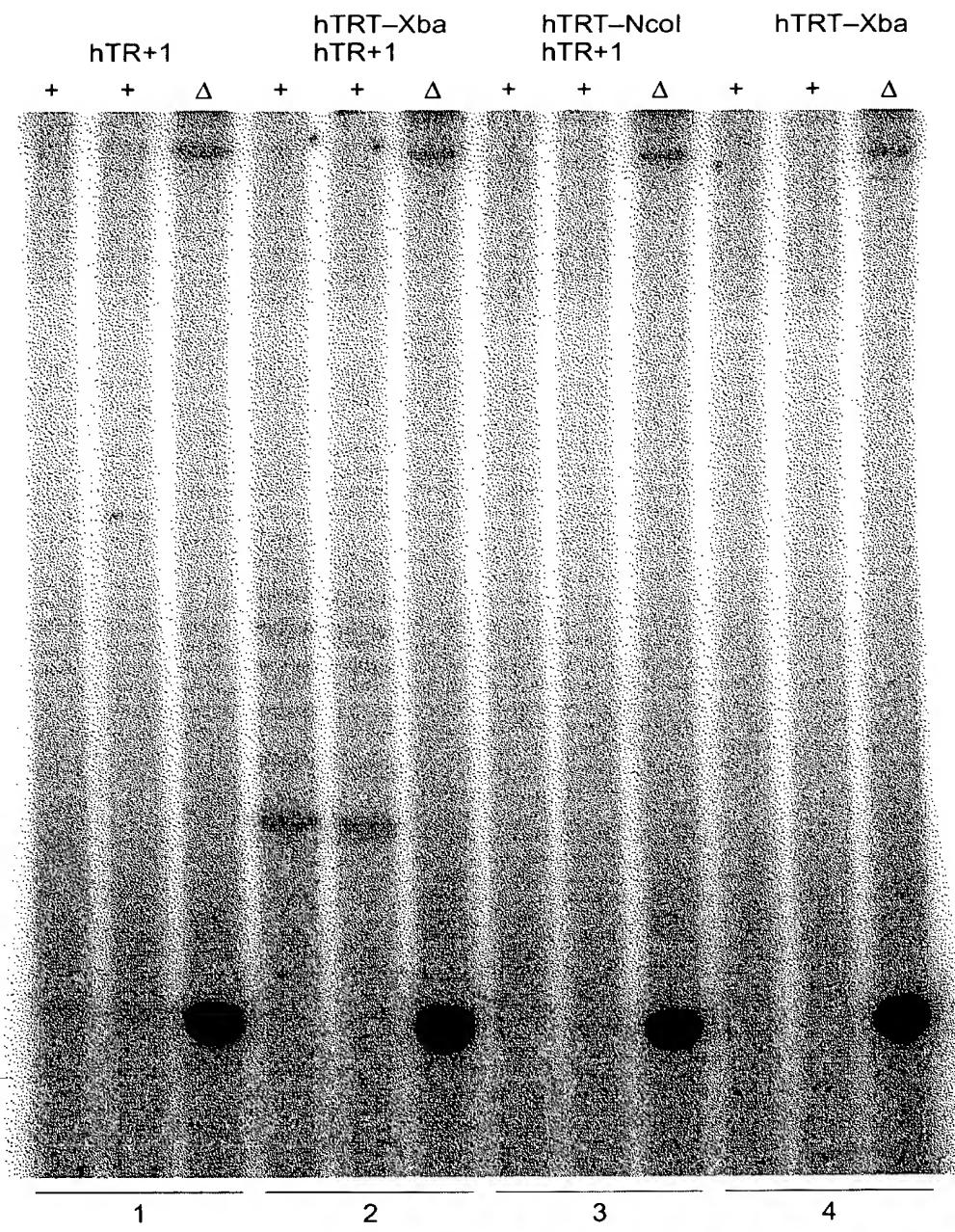


FIG. 10A

-

+

10/103

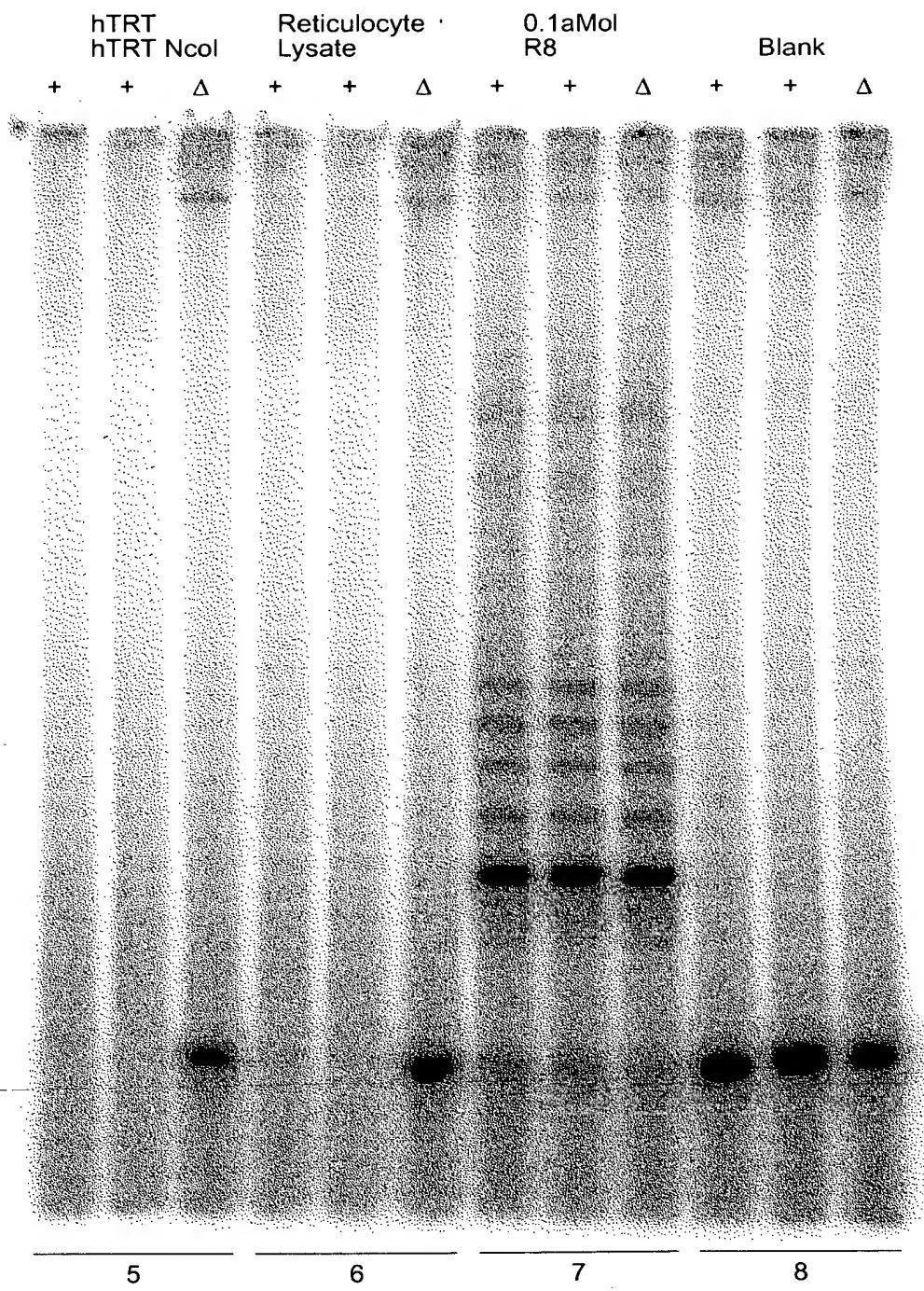


FIG. 10B

+

+

11/103

Telomerase Specific Motifs

	MOTIF T	MOTIF T'	MOTIF A	MOTIF B'	MOTIF C	MOTIF D	MOTIF E
TRT con	W1 FFY TE	Y Rk W 1 I	p 1yF D	Y q GipQGs 1s 1	111r1 DDFL it	g n K	w g s 1
hTRT	546 WLMSVYVVELLRSFFYYTETTFQKNRLLFFYRKSVWSKIQSIGI	I E V	cYD i	YYQCQQGIPQGSIIISTLLCSICLY	15 LLLRLVDDFLLLVT	15 GVPEYGCVNLRKTVV	WCGLLLDDTRTL 192
sPTRT	429 WLYNSFIIPILQSFYYTESSDLRNRTVYFRKDIIWKLLCRPFI	13 EAEVRI	69 PELYFVKVDVTGAYDTI	104 YYQCQQGIPQGSIIISTLLCSICLY	16 VLLRVVDDFLFIT	15 GFEKHNFSSTSLEKTVI	FFGGFSVNMRSL 176
Ea_p123	441 WIFEDLVVSLIRCFYYVTEQQKSYSKTYYYRKNIWIDVIMKMSI	12 ENNVR	66 RKKYFVRIDIKSICYDRI	99 YLQKVGIPQGSIISSSFLCHFYM	24 LLMRLTDYLLLT	15 VSRENGFKFNMKKLQT	WIGISIDMKT 174
Sc_Est2	366 WLFRQLIPKIIQTFFYCYTEISSTVT . IVYFRHDTWNKLLTPFI	12 EKEVE	67 PKLFFATMDIEKCYDSV	117 YKOTKGIPQGLCVSSSISSFY	18 LILKLADDFLIIS	15 GFQKYNAKANRDKILA	WKHSSTMNNNFH 141
RT con		9 ENNVC	68 PELYFMKFDVKSICYDSI	85 YIREDGGLFQGSSLSSAPIVDLY	h	h hDh AF h	hLG h
			h	GY			

Telomerase RT Motifs (Fingers)

	MOTIF 1	MOTIF 2	MOTIF A	MOTIF B'
TRT con	R iPKK	FR I	p 1yF	Y q GipQGs 1s 1
hTRT	11 SRLRFIPKPDG	0 LRPIV	69 PELYFVKVDVTGAYDTI	YYQCQQGIPQGSIIISTLLCSICLY
sPTRT	10 AVIRLLPKANT	0 FRLIT	66 RKKYFVRIDIKSICYDRI	99 YLQKVGIPQGSIISSSFLCHFYM
Ea_p123	10 GKLRLLPKKTT	0 FRPIM	67 PKLFFATMDIEKCYDSV	117 YKOTKGIPQGLCVSSSISSFY
Sc_Est2	13 SKMRIIPKKSN	2 FRIIA	68 PELYFMKFDVKSICYDSI	85 YIREDGGLFQGSSLSSAPIVDLY
RT con	p h h K	hR h	h	GY

Telomerase RT Motifs (Palm, Primer Grip)

	MOTIF C	MOTIF D	MOTIF E
TRT con	111r1 DDFL it	g	w g s 1
hTRT	15 LLLRLVDDFLLLVT	15 GVPEYGCVNLRKTVV	WCGLLLDDTRTL 192
sPTRT	16 VLLRVVDDFLFIT	15 GFEKHNFSSTSLEKTVI	FFGGFSVNMRSL 176
Ea_p123	24 LLMRLTDYLLLT	15 VSRENGFKFNMKKLQT	WIGISIDMKT 174
Sc_Est2	18 LILKLADDFLIIS	15 GFQKYNAKANRDKILA	WKHSSTMNNNFH 141
RT con	h Y DDhhh	Gh h ck h	hLG h
		F	

FIG. 11

+

+

12/103

181 GGACCCGGCGGCTTCCGCGCCTGGTGGCCCAGTCCTGGTGTGCGTGCCTGGACGC
CCTGGGCCGCGAAAGGCGCGACCACCGGGTCACGGACCACACGCACGGACCCTGCG

NFkB_CS1
GGGRQTYYQC
NFkB-MHC-I.2
TGGCCTTCCCC

241 ACGGCCGCCCCCGCCGCCCTCCTCCGCCAGGTGGCCTCCCCGGGTGGCGTCCG
TGCCGGGGGGGGCGGGAGGAAGGCGGTCCACCCGGAGGGGCCCCAGCCGCAGGC

Intron1

301 GCTGGGGTTGAGGGCGGCCGGGGGAACCAGCGACATGCCAGAGCAGCGCAGGCAGTC
CGACCCCAACTCCCGCCGGCCCCCTGGTCGCTGTACGCCCTCGTCGCGTCCGCTGAG

NFkB_CS1
GGGRQTYYQC
NFkB_CS2
RGGRMTYYCC
Topo_II_cleavage_site
RNYNNCNNGYNGKTNYNY

*****>

361 AGGGCGCTTCCCCCGCAGGTGTCTGCCTGAAGGGAGCTGGTGGCCGAGTGCTGCAGAGG
TCCCGCGAAGGGGGCGTCCACAGGACGGACTTCCTCGACCACCGGGCTCACGACGTCTCC

FIG. 12

+

1 AAAACCCCAA AACCCCAAAA CCCCTTTAG AGCCCTGCAG TTGGAAATAT
 51 AACCTCAGTA TTAATAAGCT CAGATTTAA ATATTAATTAA CAAAACCTAA
 101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC
 151 TCTTAAGACT TGTGAAGAAA TTAAAGAAC TAAAACGTTG TACTCTTGGAA
 201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAAGAT
 251 TTAGAAGATA TTAAATATT TGCGCAGAC AATATTGTTG CTACTCCACG
 301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTAA
 351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTCTT
 401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTCA
 451 ACTTAAGGGAA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTCAA
 501 CTCAAAAGCA GTATTTCTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
 551 ATTGGAAATG AGCTCTTCG ACATCTCTAC ACTAAATATT TAATATTCCA
 601 GCGAACTTCT GAAGGAACCT TTGTTCAATT TTGCGGGAA AACGTTTTG
 651 ATCATTGAA AGTCAACGAT AAGTTTGACA AAAAGAAAA AGGTGGAGCA
 701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
 751 GAATGAGAAA GATCACTTTC TCAACAAACAT CAACGTGCCG AATTGGAATA
 801 ATATGAAATC AAGAACCAGA ATATTTATT GCACTCATT TAATAGAAAT
 851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTC
 901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTAG AGATTTAATA
 951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG
 1001 CTTGAGAAAG TCAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG
 1051 TCCTCTTCCA GAAAATTGGC GGGAACGGAA AAAAAAAATC GAAAACATTGA
 1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC
 1151 TACACAACGT ATAATAAAATCGT CTCACACAA TTTATTAATG AATTTTTCTA
 1201 CAATATACCTC CAAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTC
 1251 AAAAGAAAGT TAAGAAATAT GTGGAACCTAA ACAAGCATGA ACTCATTAC
 1301 AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
 1351 GGTTGAGACC TCTGCAAAGC ATTTTTATTA TTTTGATCAC GAAAACATCT
 1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
 1451 CTGATTAGAT GATTTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
 1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
 1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
 1601 GAAGAATGGA AAAAGTCGCT TGGATTGCA CCTGGAAAAC TCAGACTAAT
 1651 ACCGAAGAAA ACTACTTCC GTCCAATTAT GACTTCAAT AAGAAGATTG
 1701 TAAATTCTAGA CCGGAAGACT ACAAAATTAA CTACAAATAC GAAGTTATTG
 1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
 1801 TTTTGGATTG GCTGTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
 1851 AGTTTGGTTG CAAATGGAAG CAAGTTGGAC AACCAAAACT CTTCTTGC
 1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAACTATC
 1951 AACATTCTCA AAAACTACTA AATTACTTTTC TTCAGATTTC TGGATTATGA
 2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTGCAAAAAC
 2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
 2101 TGCACCTGAA GGAGGACAAAT ATCCAACCTT ATTCACTGTT CTTGAAAATG
 2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAAGCAA
 2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAAGTCA TTAATATTG
 2251 CCAATATAAT TACATTAACCT TTAATGGGAA GTTTTATAAA CAAACAAAAG
 2301 GAATTCTCA AGGTCTTGA GTTTCATCAA TTTTGTCACTC ATTTTATTAT
 2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA

FIG. 13

2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
 2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
 2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
 2551 GACTAGTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
 2601 GTGTTGAGGA GCAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
 2651 TCAATTGATA TGAAAACTCT TGCTTAAATG CCAAATATTA ACTTGAGAAT
 2701 AGAAGGAATT CTGTTGACAC TCAATCTAA CATGCAAACA AAGAAAGCAT
 2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTAATGAA TAACATTACC
 2801 CATTATTTA GAAAGACGAT TACAACCGAA GACTTGCAG ATAAAACCTT
 2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
 2901 AATAACAAGGA CCACCTTAAG AAGAACTTAG CTATGAGCAG TATGATCGAC
 2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
 3001 CCTTGTGTGC AATATTAAGG ATACAATTT TGAGAGGAG CATTATCCAG
 3051 ACTTTTCCT TAGCACACTG AAGCACTTA TTGAAATATT CAGCACAAAA
 3101 AAGTACATT TCAACAGAGT TTGCATGATC CTCAGGCAA AAGAACAAA
 3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
 3201 CTATTCTAAC TTATTTGGG AAGTTAATTT TCAATTGGG TCTTATATAC
 3251 TGGGGTTTTG GGGTTTGGG GTTTGGGG

FIG. 13
(CONTINUED)

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWI0KVIRC RNQSQSHYKD
 51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
 101 SSSDVSDRQK LQCFGQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
 151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKQKQGGA
 201 ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNNMKSRTT IFYCHFNRN
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAYM
 301 LEVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHели
 401 KNLLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKL RWIFEDLVVS
 451 LIRCFFYVTE QQKSYSKTYY YRKNIWDVIM KMSIADLKKE TLAEVQEKEV
 501 EEWKKSLGFA PGKLRLIPKK TTFRPIMTFN KKIVNSDRKT TKLTNTKLL
 551 NSHMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN
 651 FRKEMKDVF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ
 701 RNYFKKDNLQ QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSIILSSFY
 751 ATLEESSLGF LRDESMNPEN PVNVNLLMRLT DDYLLITTQE NNAVLFIEKL
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
 851 SIDMKTLALM PNINLRIEGI LCTLNLMQTKKASMWLKKK LKSFLMNNIT
 901 HYFRKTITTE DFANKTLNKL FISGGYKYMQ CAKEYKDHFK KNLAMSSMID
 951 LEVSKIIYSV TRAFFKYLVN NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK
 1001 KYIFNRVCMI LKAKEAKLKS DQCQSLIQYD A

FIG. 14

15/103

+

FIG. 15

-

16/103

FIG. 15
(CONTINUED)

+

17/103

2268	TTT	GAG	ATA	ATA	TTA	AAA	G	gtattgtataaaattttaccacttaacgattttaccag	AC	CTC	GAA	ACT	2336									
366	F	E	I	I	L	K	D		L	E	T		375									
2337	TTC	TTG	AAA	TTA	TCG	AGA	TAC	GAG	TCT	TTT	AGT	TTA	ATG	AGT	AAC	ATA	AAG	2396				
376	F	L	K	L	S	R	Y	E	S	F	S	L	H	Y	L	M	S	N	I	K	395	
2397	gtaatatgccaaatttttaccattaaacaatcg	ATT	TCA	GAA	ATT	GAA	TGG	CTA	GTC	CTT	GGA										2465	
396	I	S	E	I	E	W	L	V	L	G										405		
2466	AAA	AGG	TCA	AAT	GCG	AAA	ATG	TGC	TTA	AGT	GAT	TTT	GAG	AAA	CGC	AAG	CAA	ATA	TTT	GCG	2525	
406	K	R	S	N	A	K	M	C	L	S	D	F	E	K	R	K	Q	I	.F	A	425	
2526	GAA	TTC	ATC	TAC	TGG	CTA	TAC	AAT	TCG	TTT	ATA	ATA	CCT	ATT	TTA	CAA	TCT	TTT	TAT		2585	
426	E	F	I	Y	W	L	Y	N	S	F	I	I	P	I	L	Q	S	F	F	Y	445	
2586	ATC	ACT	GAA	TCA	AGT	GAT	TTA	CGA	AAT	CGA	ACT	GTT	TAT	TTT	AGA	AAA	GAT	ATT	TTT	TAT		2645
446	I	T	E	S	S	D	L	R	N	R	T	V	Y	F	R	K	D	I	W	K	465	
2646	CTC	TTG	TGC	CGA	CCC	TTT	ATT	ACA	TCA	ATG	AAA	ATG	GAA	GCG	TTT	GAA	AAA	ATA	AAC	GAG	2705	
466	L	L	C	R	P	F	I	T	S	M	K	M	E	A	F	E	K	I	N	E	485	
2706	gtattttaaagtatttttgcaaaaaggctaataattttcag	AAC	AAT	GT	AGG	ATG	GAT	ACT	CAG	AAA	ACT	2775										
486	N	N	V	R	M	D	T	Q	K	K	T										495	
2776	ACT	TTG	CCT	CCA	GCA	GTT	ATT	CGT	CTA	TTA	CCT	AAG	AAG	AAT	ACC	TTT	CGR	CTC	ATT	ACG	2835	
496	T	L	P	P	A	V	I	R	L	L	P	K	K	N	T	F	R	L	I	T	515	
2836	AAT	TTA	AGA	AAA	AGA	TTC	TTA	ATA	AAG	gtatttaatttttggcatcaatgtactttactttcaatcttata											2906	
516	N	L	R	K	R	F	L	I	K												524	
2907	ttagcag	ATG	GGT	TCA	AAC	AAA	ATG	TTA	GTC	AGT	ACG	AAA	ACT	TTA	CGA	CCT	GTG				2967	
525	M	G	S	N	K	K	M	L	V	S	T	N	Q	T	L	R	P	V		542		
2968	GCA	TCG	AAA	CAT	TTA	ATC	AAA	GAA	GAA	AGT	AGT	GGT	ATT	CCA	TTT	AAC	TTG	GAG		3027		
543	A	S	I	L	K	H	L	I	N	E	E	S	S	G	I	P	F	N	L	E	562	
3028	GTT	TAC	ATG	AAG	CTT	CTT	ACT	TTT	AAG	AAG	GAT	CTT	CTT	AAG	CAC	CGA	ATG	TTT	GG	gtaat	3088	
563	V	Y	M	K	L	L	T	F	K	K	D	L	L	K	H	R	M	F	G	581		

FIG. 15
(CONTINUED)

+

+

18/103

3089 tataataatgcgcgattccatttttcagg G CGT AAG AAG TAT TTT GTC CCG ATA GAT ATA
 582 R K K Y F V R I D I 3155
 591
 3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC
 592 K S C Y D R I K Q D L M F R I V K K K L 3215
 611
 3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT
 612 K D P E F V I R K Y A T I H A T S D R A 3275
 631
 3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagtttattttcatggaaatttttaacaa 3343
 632 T K N F V S E A F S Y F 643
 3344 attcttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA
 644 D M V P F E K V V Q L L S M K T 3405
 659
 3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT
 660 S D T L F V D Y W T K S S E I F 3465
 679
 3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaaattgtgaatttaacaa 3532
 680 K M L K E H L S G H I V K 692
 3533 ctaatggaaacttag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA
 693 I G N S Q Y L Q K V G I P Q G S 3593
 708
 3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG
 709 I L S F L C H F Y M E D L I D E Y L S 3653
 728
 3654 TTT ACG AAA AAG GAA TCA GTG TTG TTA CGA GTC GAT TTC CTC TTT ATA ACA
 729 F T K K G S V L L R V V D D F L P I T 3713
 748
 3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgtcattcc 3777
 749 V N K K D A K F L N L S L R G 764
 765
 3778 taagtcttaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA
 778
 3840 F E K H N F S T S L E K T V 3840
 779 I N F E N S N G I I N N T F F N E S K K 798 3900
 798

FIG. 15
(CONTINUED)

+

19/103

FIG. 15
(CONTINUED)

FIG. 15
(CONTINUED)

1 gcagcgctgc gtcctgctgc gcacgtggga agccctggcc ccggccaccc cgcgatgcc
 61 gcgcgctccc cgctgcccggag ccgtgcgctc cctgctgcgc agccactacc gcgagggtgt
 121 gccgctggcc acgttctgtgc ggccgcctggg gccccagggc tggcgctgg tgcagcgccg
 181 ggacccggcg gcttcccgcg cgctgggtgc ccagtgcctg gtgtgcgtgc cctggacgc
 241 acggccgccc cccgcccggcc cctcctccg ccaggtgtcc tgcctgaagg agctgggtggc
 301 ccgagtgtcg cagaggctgt gcgagcgcgg cgcaagaacat gtgtggct tcggcttcgc
 361 gtcgtggac gggggcccg gggggcccccc cgaggccctc accaccaggcg tgccgagcta
 421 ctcgtggccaa acggtgacccg acgactcgcc ggggagcggg gctgtgggc tgctgtgc
 481 ccgcgtggc gacgacgtgc ttgttcaacct gtcggcacgc tgccgcgtct ttgtgtgt
 541 ggctcccgac tgcgcctacc aggtgtgcgg gccgcgcgtg taccagctcg gcgctgccac
 601 tcaggccccgg cccccggccac acgcttagtgg accccgaagg cgtctggat gcgaaacgggc
 661 ctggaaaccat aegtcgtacgg aggccggggg cccctgtggc ctgcacggccc cgggtgcgag
 721 gaggcgccgg ggcagtgcga gccgcgtgtc gccgttgcggc aagaggccca ggcgtggcgc
 781 tgcccttgag ccggagccgg cgcgggttgg gcaagggttc tggccacc cggcaggac
 841 gctgggaccg agtgcacgtg gtttctgtgt ggtgtcacct gccagaccgg ccaagaagac
 901 caccttttgc gagggtgcgc tctctggcac gcgcactcc caccatccg tggccgcca
 961 gcaccacgcg ggccccccat ccacatcgcg gcaccacacgt ccctggaca cgccttgtcc
 1021 cccgggtgtac gccgagacca agcacttcct acttcctca ggcgcacaagg agcagctgc
 1081 gccctcttc ctactcgat ctctgaggcc cagctgtact ggccgtcgga ggctcgta
 1141 gaccatctt ctgggttcca ggcctggat gccaggact ccccgccagg tgccctggct
 1201 gccccagcgc tactggcaaa tgccggccct gtttctggag ctgttggga accacgcga
 1261 gtggccctac ggggtgtcc tcaagacgcg ctgcggcgtg cgagctgcgg tcacccca
 1321 agccgggtgtc tggcccccgg agaagccca ggctctgtg ggcggccccc aggaggagg
 1381 cacagacccc cgtcgccctgg tgcaaggtgtc ccgcgcacgc agcagccct ggcagggtgt
 1441 cggcttcgtg cgggcctgcg tgccggcgt ggtgcggccca ggccttggg gctccaggca
 1501 caacgaacgc cgcttcctca ggaacaccaa gaagttcatc tccctggga agcatgcca
 1561 gtcctcgctg caggagctga cgtggaaagat gagcgtgcgg gactgcgtt ggctgcgc
 1621 gagcccaagg gttggctgtg ttccggccgc agagcaccgt ctgcgtgagg agatctggc
 1681 caagttctt cactggctga tgagtgtga ctcgtcgag ctgcgttggg ctttcttt
 1741 tgtcacggag accacgtttc aaaagaacag gtccttttc taccggaaa gtgtctggag
 1801 caagttgcaaa agcattggaa tcagacacga cttaaagggg gtcagactgc gggagctgtc
 1861 ggaaggcagag gtcaggcgcg atcggggaaac caggccgcg ctgcgtacgt ccagactcc
 1921 cttcatcccc aagcctgacg ggctggccg gattgtgaac atggactacg tcgtggagc
 1981 cagaacgttc cgcagagaaa agagggccga gcgttcaccc tcgagggtga aggcaactgtt
 2041 cgcgtgtcc aactacgacg gggcggccg cccggccctc ctggccgcct ctgtctggg
 2101 ctcggacat attcacaggg ctcggccac ctgcgtgtc ctgcgtcgcc cccaggaccc
 2161 gccgcctgag ctgtactttg tcaagggtgg ttcgtacggc gctgtacgaca ccatccccca
 2221 ggacaggctc acggagggtca tcgcccggcat catcaaacc cagaacacgt actgcgtgc
 2281 tcggtatgcc gttggccaga aggccggccca tgggcacgtc cgcaaggccct tcaagagcca
 2341 cgtcttacc ttgacagacc tccagccgtc catgcacacg ttctggctc acctgcaggaa
 2401 gaccaggccg ctgagggtat cgcgtgtcat cgcacggc tccctgtca atgaggcc
 2461 cagtggcctc ttgcacgtct tcctacgctt catgtggccac cacccgtgc gcatcagg
 2521 caagtcttac gtcacgtgtc agggatccc gcagggtctc atccttcca cgctgtctg
 2581 cagcctgtgc tacggcgaca tggagaacaa gtcgtttcg gggattccgc gggacgggct
 2641 gtcctgtcg tgggtggatg atttcgtt ggtgacaccc caccctaccc acgcgaaaac
 2701 ctccttcagg acctgtgtc gagggtccc tggatgtggc tgcgtgtga acttgcggaa
 2761 gacagtggta aacttccctg tagaagacga gggccctgggt ggcacggctt ttgttc
 2821 gcccggccac ggcttattcc cttgggtgcgg ctcgtgtctg gataccggg ccctggaggt
 2881 gcagagcgac tactccagct atgcccggac ctccatcaga gccagtc tca cttcaacc
 2941 cggcttcaag gtcgggagga acatgcgtcg caaatcttt ggggtcttc ggtgtaaag
 3001 tcacagcctg ttctgtgtt tgcagggtaa cagccctccag acgggtgtca ccaacatcta
 3061 caagatctc ctgtcgagg cgtacaggat tcaacgtat gtgtcgacg tcccatttca
 3121 tcacgttcaat tggaaaacc ccacatccc ctgcgcgtc atctctgaca cggccctcc
 3181 ctgtacttcc atcctgtaaag ccaagaacgc agggatgtcg ctggggccca agggcgcc
 3241 cggccctctg ccctccggagg cctgtcgatc gtcgtgtcc acgcatcc tgctcaag
 3301 gactcgacac cgtgtcaccc acgtgcaccc ctgggggtca ctgcggacag cccagacgc
 3361 gtcgtgtcg aagctcccg ggacgcgtc gtcgtccctg gggccgcag ccaaccggc
 3421 actgccttca gacttcaaga ccacatccc gtcgtgtcc acgtgcaccc cccggccaca
 3481 gagcagacac cagcagccct gtcacggccg gtcgtgtcc ccaggagggg aggggggg
 3541 cacacccagg cccgcacccgc tggggatgtc aggctgtat ggtgtttgg cggaggcc
 3601 catgtccggc tgaaggctga gtgtccggc gaggctgtgg gtcgtgtcc gccaagg
 3661 gagtgtccag cacacctgcg gtgttcaactt cccacaggc tggcgctcg ctccaccc
 3721 gggccagctt ttcttcacca ggagccggc ttccactccc cacaatggaa tagtccat
 3781 ccacatttcg cattgttccat ccctcgccctt gtcgtgtcc acgttccacc cccaccatcc
 3841 aggtggagac cctgagaagg accctggag ctctggaaat ttggagtgac caaagg
 3901 ccctgtacac aggcgaggac ctcgcacccg gatgggggtc ctcgtgggtc aaattgggg
 3961 gaggtgtgt gggagtaaaa tactgaatat atgagtttt cagtttgaa aaaaa

FIG. 16

22/103

MPRAPRCAVRSLLRSHYREVPLATFVRRLLGPQGWRLVQRGDP
AAFRALVAQCLVCVPWDARPPAAPSFRQVSCLKEVARVLQRL
CERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPNTVTDALR
GSGAWGLLRLRVGDVLVHLLARCALFVLPAPSCAQVCGPPLY
QLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPG
ARRRGGSASRSLPLPKRPRRGAAPEPERTPVQGWSWAHPGRTRG
PSDRGFCVVS PARPAAEATSLEGALSGTRHSHPGSVRQHHAGPP
STSRPPRPWDTCPVYAEKHFHFLYSSGDKEQLRPSFLLSLRP
SLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRLFLLEL
LGNHAQCPTYGVLLKTHCPLRAVTPAAGVCAREKPQGSVAAPEE
EDTDPRRLVQLLRQHSSPWQVYGFVRACLRRLVPPGLWGSRHNE
RRFLRNTKKFISLGKHAKLSLQELETWKMSVRDCAWLRRSPGVGC
VPAAEHRLREELAKFLHWLMSVYVVELRSFFYVTETTFQKNR
LFFYRKSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREARPAL
LTSRRLRFIPKPDGLRPIVNMDDYVVGARTFRREKRAERLTSRVKA
LFSQLNRYEARRPGLLGASVGLLDDIHRAWRTFVLRVRAQDPPP
ELYFVKVDVTGAYDTIPQDRLTEVIASI1KPQNTYCVRRYAVVQ
KAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI
EQSSSLNEASSGLFDVFLRFMCHAVRIRGKSYSVQCQGI PQGSI
LSTLLCSLCYGDMEKLFAGIRRDLQVNSLQTVCNTNIYKILLQAYRF
KTFLRTLTVRGPYEYGCVVNLRKTVNF PVEDEALGTAFVQMPA
HGLFPWCMLLDDRTLEVQSDYSSYARTSIRSALTNRGFKAGR
NMRRKLFVLRLLKCHSFLDLQVNSLQTVCNTNIYKILLQAYRF
HACVLQLPFHQVWKNPTFFLVRVISDTSALCYSILKAKNAGMSL
GAKGAAGPLPSEAVOWLCHQAFLKLTRHRTVYVPLLGSRLTAQ
TQLSRKLPGTTLTALEAAAANPALPSDFKTIID

FIG. 17

GGCCAAGTCTCTGCACTGGCTGATGAGTGTGACGTCGAGCTGCTCAGGTCTTTCTT
TTATGTCACGGAGACCACGTTCAAAAGAACAGGCTTTCTACCGGAAGAGTGTCTG
GAGCAAGTGTCAAAGCATTGAATCAGACAGCACTGAAGAGGGTGCAGCTCGGGAGCT
GTCGGAAGCAGAGGTCAAGCAGCATGGGAAGGCCAGGCCCTGCTGACCTCCAGACT
CCGCTTATCCCCAAGCCTGAGGGCTCGGGCGATGTGAACATGGACTACGTCGTTGG
AGCCAGAACGCTTCCGAGAAAAAGAGGGCGAGCGCTCCTACCTCGAGGGTGAAGGCACT
GTTTCAGCGTGCTCAACTACGAGCGGGCGCGCCGCTCTGGCGCCTCTGTGCT
GGGCCTGGACGATATCCACAGGGCTGGCGCACCTCGTGTGCGTGTGCGGGCCAGGA
CCCGCCGCTGAGCTGACTTGTCAAGGGATGTGACGGGCGTACGACACCATCCC
CCAGGACAGGCTCACGGAGGTATGCCAGCATCATCAAACCCAGAACACGTAUTGCGT
GCGTCGGTATGCCGTGGTCCAGAAGGCCCGCCATGGCACGTCGCAAGGCCCTCAAGAG
CCACGCTCTACGTCAGTGCAGGGGATCCCGCAGGGCTCCATCTCCACGCTGCTCT
GCAGCCTGTGCTACGGCAGATGGGAACAAAGCTTTGCGGGGATTGGCGGGACGGC
TGCTCCTCGTGTGGATGATTCTTGTGGTGTACACCTCACCTCACCCACGCAAAA
CCTTCCTCAGGACCTGGTCCAGGTGTCCTGAGTATGGCTCGTGGTGAACCTTGCGGA
AGACAGTGGTGAACCTCCCTGAGAAGACGAGGCCCTGGTGGCACGGCTTTGTTCAAGA
TGCCGGCCACGGCTATTCCCTGGTGCAGGCCCTGCTGGATAACCGGACCTGGAGG
TGCAGAGCAGACTACCCAGTATGCCGACCTCCATCAGAGCCAGTCTCACCTCAACC
GCGGCTCAAGGCTGGAGGAACATGCGTCGCAAACATCTTGGGTTCTGCGGCTGAAGT
GTCACAGCCTTCTGGATTGCAAGGTGAACAGCCTCCAGACGGTGTGACCAACATCT
ACAAGATCCTCTGTCGAGGGTCAAGGTTTACCGCATGTTGCTGCTGAGCTCCCATTT
ATCAGCAAGTTGGAGAACCCACATTTCCTGCGCTCATCTGACACGGCCTCCC
TCTGCTACTCCATCCTGAAAGCAAGAACGCAAGGGATGTCGCTGGGGCAAGGGCGCG
-CCGGCC7TCTGCCCTCCAGGGCGTGCAGTGGCTGCCACCAAGCATTCTGCTCAAGC
TGACTCGACACCGTGTCACTACGTCGCACTCCCTGGGTCACTCAGGACAGCCCAGACGC
AGCTGAGTCGGAAGCTCCGGGACGACGCTGACTGCCCTGGAGGCCGAGCCAACCCGG
CACTGCCCTCAGACCTCAAGACCATCTGGACTGATGGGCCACCGGCCACAGCCAGGGCG
AGAGCAGACACCGACAGGGCTTCAGCAGGGCTCACTGTCAGTCCCAGGGAGGGAGGGCG
CCACACCCAGGCCCTGCAACCGCTGGGAGTGTGAGGCCCTGAGGCTGAGCAGTGTCCAGGCCAAGGGC
GCATGTCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCCTGAGGCCAGTGTCCAGGCCAAGGGC
TGAGTGTCCAGCACACCTGCCCTTCACCTCCCCACAGGCTGGCGCTGGCTCCACCCC
AGGGCCAGCTTCTCACCAAGGAGGCCGGCTTCACTCCCCACATAGGAATAGTCCATC
CCCAGATTGCCATGTTCACCCCTGCCCTGCCCTCTTGCCTCCACCCCCACCATC
CAGGGTGGAGAACCTGAGAAGGACCCCTGGGAGCTCTGGGAATTGGAGTGCACCAAGGGT
GCCCTGTACACAGGGAGGGACCTGCACTGGATGGGGTCCCTGTGGGTCAAATTGGGG
GGAGGTGCTGGGGAGTAAAATACTGAATATGAGTTTCACTGAGTTTCACTGAGTTTCA
AAAAAAAAAAAAAA

FIG. 18

+

23/103

Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe Phe
 Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe
 Phe Tyr Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile
 Gly Ile Arg Gln His Leu Lys Arg Val Gln Leu Arg Glu
 Leu Ser Glu Ala Glu Val Arg Gln His Arg Glu Ala Arg
 Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile Pro Lys
 Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val
 Val Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu
 Arg Leu Thr Ser Arg Val Lys Ala Leu Phe Ser Val Leu
 Asn Tyr Glu Arg Ala Arg Arg Pro Gly Leu Leu Gly Ala
 Ser Val Leu Gly Leu Asp Asp Ile His Arg Ala Trp Arg
 Thr Phe Val Leu Arg Val Arg Ala Gln Asp Pro Pro Pro
 Glu Leu Tyr Phe Val Lys Val Asp Val Thr Gly Ala Tyr
 Asp Thr Ile Pro Gln Asp Arg Leu Thr Glu Val Ile Ala
 Ser Ile Ile Lys Pro Gln Asn Thr Tyr Cys Val Arg Arg
 Tyr Ala Val Val Glu Lys Ala Ala His Gly His Val Arg
 Lys Ala Phe Lys Ser His Val Leu Arg Pro Val Pro Gly
 Asp Pro Ala Gly Leu His Pro Leu His Ala Ala Leu Gln
 Pro Val Leu Arg Arg His Gly Glu Gln Ala Val Cys Gly
 Asp Ser Ala Gly Arg Ala Ala Pro Ala Phe Gly Gly

FIG. 19

1
met

GCAGCGCTGCGTCCTGCTGCGCACGTGGAAAGCCCTGGCCCCGGCCACCCCCGGCG ATG

10

pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
 CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

20

his tyr arg glu val leu pro leu ala thr phe val arg arg leu
 CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

30

gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
 GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

40

phe arg ala leu val ala gln cys leu val cys val pro trp asp
 TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

50

ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
 GCA CGG CCG CCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

60

leu lys glu leu val ala arg val leu gln arg leu cys glu arg
 CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

70

gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
 GGC GCG AAG AAC GTG CTG GCC TTC GGC TTG CGC CTG GAC GGG

80

ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
 GCC CGC GGG GGC CCC CCC GAG GGC TTC ACC ACC AGC GTG CGC AGC

90

100

110

120

FIG. 20

+

+

24/103

130

tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

140

trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

160

leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

170

ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

190

thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

200

leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

220

val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTC CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

230

ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

250

ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

260

ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

280

val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

290

gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

310

gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

320

trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GGC GAG ACC AAG CAC TTC

330

FIG. 20
(CONTINUED)

4

+

340

leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
 CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

350

leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
 CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

360

glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
 GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

370

arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
 CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

380

leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
 CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

390

val leu leu lys thr his cys pro leu arg ala ala val thr pro
 GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

400

ala ala gly val cys ala arg glu lys pro gln gly ser val ala
 GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

410

ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
 GCC CCC GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

420

leu arg gln his ser ser pro trp gln val tyr gly phe val arg
 CTC CGC CAG CAC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

430

ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
 GCC TGC CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

440

his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
 CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

450

leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
 CTG GGG AAG CAT GCC AAG CTC TCG CAG GAG CTG ACG TGG AAG

460

met ser val arg asp cys ala trp leu arg arg ser pro gly val
 ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

470

gly cys val pro ala ala glu his arg leu arg glu glu ile leu
 GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

480

510

490

520

530

540

FIG. 20
(CONTINUED)

550

ala lys phe leu his trp leu met ser val tyr val val glu leu
 GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

560

leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
 CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

580

arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser
 AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC

590

ile gly ile arg gln his leu lys arg val gln leu arg glu leu
 ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

610

ser glu ala glu val arg gln his arg glu ala arg pro ala leu
 TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

620

leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
 CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

640

pro ile val asn met asp tyr val val gly ala arg thr phe arg
 CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

650

arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
 AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

660

phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
 TTC AGC GTG CTC AAC TAC GAG CGG GCG CGC CCC GGC CTC CTG

670

gly ala ser val leu gly leu asp asp ile his arg ala trp arg
 GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

680

thr phe val leu arg val arg ala gln asp pro pro pro glu leu
 ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG

690

tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
 TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

700

710

720

730

740

750

gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
 CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
 AAC ACG TAC TGC GTG CGT TAT GCC GTG GTC CAG AAG GCC GCC

FIG. 20
(CONTINUED)

+

27/103

760

his gly his val arg lys ala phe lys ser his val leu arg pro
 CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA

770

val pro gly asp pro ala gly leu his pro leu his ala ala leu
 GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG

780

gln pro val leu arg arg his gly glu gln ala val cys gly asp
 CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT

790

ser ala gly arg ala ala pro ala phe gly gly OP
 TCG GCG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTGGT

GACACCTCACCTCACCCACGCGAAAACCTTCCTCAGGACCCCTGGTCCGAGGTGTCCCTGA

GTATGGCTGCGTGGTGAACCTGCGGAAGACAGTGGTGAACCTCCCTGTAGAACAGCAGGCC

CCTGGGTGGCACGGCTTGTTCAGATGCCGGCCACGGCCTATTCCCCTGGTGCAGGCC

GCTGCTGGATAACCGGACCCCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCGACCTC

CATCAGAGCCAGTCTCACCTCAACCGCGGCTTCAAGGCTGGAGGAACATGCGTCGCAA

ACTCTTTGGGTCTTGCCTGAAAGTGTACAGCCTGTTCTGGATTTGCAGGTGAACAG

CCTCCAGACGGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTCA

CGCATGTGTGCTGCAGCTCCATTCTACAGCAAGTTGAAGAACCCACATTTCCT

CGCGTCATCTCTGACACGGCCTCCCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG

GATGTCGCTGGGGCCAAGGGCGCCGCCCTGCCCCCTGAGGCGTGCAGTGGCT

GTGCCACCAAGCATTCTGCTCAAGCTGACTCGACACCGTGTACCTACGTGCCACTCCT

GGGGTCACTCAGGACAGCCCAGACGCGACTGAGTCGAAGCTCCGGGACGACGCTGAC

TGCCCTGGAGGCCGAGCCAACCCGGCACTGCCCTCAGACTCAAGACCACCTGGACTG

ATGGCCACCCGCCACAGCCAGGCCAGAGCAGACACCAGCAGGCCCTGTCACGCCGGCT

CTACGTCCCAGGGAGGGAGGGCGGCCACACCCAGGCCAGCGCTGGGAGTCTGAGG

CCTGAGTGAGTGTGTTGGCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG

GCCTGAGCGAGTGTCCAGCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTCACCTCCC

CACAGGCTGGCGCTGGCTCCACCCAGGGCAGCTTTCCTCACCGAGGAGGCCGCTTC

CACTCCCCACATAGGAATAGTCCATCCCCAGATTGCCATTGTTCACCCCTGCCCTGCC

CTCCTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCCTGAGAAGGACCCCTGGGAGCTC

TGGGAATTGGAGTGACCAAGGTGTGCCCTGTACACAGGGCGAGGACCCCTGCACCTGGAT

GGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATG

AGTTTTTCAGTTGAAAAAAAAAAAAAA

FIG. 20
(CONTINUED)

+

1 CCATGGGACCCACTGCAGGGCAGCTGGGAGGCTGCAGGCTTCAGGTCCCAGTGGGGTTG
GGTACCCCTGGGTGACGTCCCCGTCACCCTCCGACGTCCGAAGTCCAGGGTCACCCCAAC

61 CCATCTGCCAGTAGAAACCTGATGTAGAATCAGGGCGCAGTGTGGACACTGTCTGAAT
GGTAGACGGTCATCTTGGACTACATCTTAGTCCCGCCTCACACCTGTGACAGGACTTA

121 CTCAAATGTCTCAGTGTGCTGAAACATGTAGAAATTAAAGTCCATCCCTCTACTCTAC
GAGTTACAGAGTCACACACGACTTTGTACATCTTAATTTCAGGTAGGGAGGATGAGATG

181 TGGGATTGAGCCCTTCCCTATCCCCCCCCAGGGCAGAGGGAGTTCCCTCTCACTCCTGTG
ACCCTAACACTCGGGGAAGGGATAGGGGGGGTCCCCGTCTCCCTCAAGGAGAGTGAGGACAC

241 GAGGAAGGAATGATACTTTGTTATTTTCACTGCTGGTACTGAATCCACTGTTCATTTG
CTCCTCCTTACTATGAAACAATAAAAGTGACGACCATGACTTAGGTGACAAAGTAAAC

301 TTGGTTTGTGTTTGTGTTGAGAGGCAGGTTCACTCTGTTGCTCAGGCTGGAGGGAG
AACCAAACAAACAAACAAACTCTCCGCCAACGTGAGAACAAACGAGTCCGACCTCCCTC

361 TGCAATGGCGCATCTGGCTACTGCAGCCTGCCTCCAGGTTCAAGTGATTCTCCT
ACGTTACCGCGCTAGAACCGAATGACGTGGAGACGGAGGGTCCAAGTCAACTAAGAGGA

alu

421 GCTTCCGCCTCCATTGGCTGGATTACAGGCACCCGCCACCATGCCAGCTAATTTT
CGAAGGCGGAGGGTAAACCGACCTAATGTCGTGGCGGTGGTACGGTCGATTAAAAA

==

481 TGTATTTTAGTAGAGACGGGGTGGGGTGGGTTCACCATGTTGCCAGGCTGGTCTC
ACATAAAATCATCTGCCACCCACCCCAAGTGGTACAACGGTCCGACCAGAG

CAP

=====>

541 GAACTCTGACCTCAGATGATCCACCTGCCTCTGCCCTCTAAAGTGTGGATTACAGGT
CTTGAAGACTGGAGTCTACTAGGTGGACGGAGACGGAGGATTCAAGACCCCTAATGTCCA

601 GTGAGCCACCATGCCAGGTCAAGATTACTCTGTTAGAAACATCTGGGTCTGAGGTAG
CACTCGGTGGTACGGTCGAGTCTTAAATGAGACAAATCTTGTAGACCCAGACTCCATC

CCAAT

*****>

661 GAAGCTCACCCACTCAAGTGTGGTGTGTTAAGCCAATGATAGAATTTTTATTGT
CTTCGAGTGGGTGAGTTCACACACCAAAATCGGTTACTATCTTAAAAAATAACA

721 TGTTAGAACACTCTTGATGTTTACACTGTGATGACTAAGACATCATCAGCTTTCAAAG
ACAATCTTGTGAGAACTACAAATGTGACACTACTGATTCTGTAGTCGAAAAGTTTC

FIG. 21

+

CAP
*****>

781 ACACACTAACTGCACCCATAATACTGGGTGTCTCTGGTATCAGCGATCTTCATTGAA
TGTGTGATTGACGTGGTATTATGACCCCACAGAAGACCCATAGTCGCTAGAAGTAACCT

CAP
*****>

841 TGCCGGGAGGCCTTCCTGCCATGCACATGGTGTAAATTACTCCAGCATAATCTCTGC
ACGGCCCTCGCAAAGGAGCGGTACGTGTACCAACATAATGAGGTCGTATTAGAACG

***>

901 TTCCATTCTTCTTCCCTCTTTAAAATTGTGTTTCTATGTTGGCTCTGCAGAG
AAGGTAAAGAAGAGAAGGGAGAAATTAAACACAAAAGATAACACCGAAGAGACGTCTC

CAP
*****>

961 AACCAAGTGTAAGCTACAACCTAACCTTGTGGAACAAATTTCACCGCCCCTTGC
TTGGTCACATTGATGTTGAATTGAAAACAACCTTGTAAAGGTTGGCGGGAAACG

1021 CCTAGTGGCAGAGACAATTACAACACAGCCCTTAAAAGGCTTAGGGATCACTAAGG
GGATCACCGTCTCTGTTAAGTGTGTCGGAAATTTCGAATCCCTAGTGATTCC

1081 GGATTCTAGAAGAGCGACCCGTAATCCTTAAGTATTACAAGACGAGGCTAACCTCCAG
CCTAAAGATCTTCTCGCTGGCATTAGGAATTCAATAATGTTCTGCTCCGATTGGAGGTC

1141 CGAGCGTGACAGCCCAGGGAGGGTGCAGGGCTGTTCAAATGCTAACGCTTCCATAAATAA
GCTCGCACTGTCGGTCCCTCCCACGCTCCGGACAAGTTACGATTGAAGGTATTATT

1201 AGCAAATTTCCTCCGGCAGTTCTGGAAAGTAGGAAAGGTTAACATTAAAGGTTGCGTTT
TCGTTAAAGGAGGCCGTCAAAGACCTTCATCCTTCAAATTGTAATTCCAACGCAA

1261 GTTAGCATTTCAGTGTGTTGCCGACCTCAGCTAACAGCATCCCTGCAAGGCCTGGAGAC
CAATCGTAAAGTCACAAACGGCTGGAGTCGATTGTCGTAGGGACGTTCCGGAGCCCTCTG

1321 CCAGAAAGTTCTGCCCTTAGATCAAACCTGAGCAACCCGGAGTCTGGATTCTGGGAA
GGTCTCAAAGAGCGGGGAATCTAGGTTGAACTCGTTGGCCTCAGACCTAAGGACCCT

TopoII
*****>

1381 AGTCCTCAGCTGTCCTGCGGTTGTGCGGGGCCCCAGGTCTGGAGGGACCAGTGGCCGT
TCAGGAGTCGACAGGACGCCAACACGGCCCCGGGTCCAGACCTCCCTGGTCACCGGCA

1441 GTGGCTCTACTGCTGGCTGGAAGTCGGCCTCTAGCTCTGCAGTCGAGGCTGGAG
CACCGAAGATGACGACCCGACCTCAGCCGGAGGATCGAGACGTCAGGCTCCGAACCTC

1501 CCAGGGCCTGGACCCCGAGGCTGCCCTCCACCCGTGCGGGCGGGATGTGACCAGATGT
GGTCCACGGACCTGGGGCTCCGACGGAGGTGGACACGCCCGCCCTACACTGGTCTACA

1561 TGGCCTCATCTGCCAGACAGAGTGCCTGGGCCCCAGGTCAAGGCCGTTGTGGCTGGTGTG
ACCGGAGTAGACGGTCTGTCACGGCCCCGGTCCAGTTCCGGCAACACCGACCACAC

1621 AGGCGCCCGGTGCGCGGCCAGCAGGAGCGCCTGGCTCCATTCCCACCCCTTCGACGG
TCCGCGGGCACGCGCCGGTCTCGCAGGACCGAGGTAAAGGGTGGAAAGAGCTGCC

FIG. 21
(CONTINUED)

+

30/103

1681 GACCGCCCCGGTGGGTGATTAACAGATATTGGGGTGGTTGCTCATGGTGGGGACCCCT
CTGGGGGCCACCCACTAATTGTCTATAACCCCACCAAACGAGTACCACCCCTGGGGAA

1741 CGCCGCCTGAGAACCTGCAAAGAGAAATGACGGGCCTGTGTCAAGGAGCCAAAGTCGCGG
GCAGGGACTCTTGACGTTCTTACTGCCGGACACAGTCCTCGGTTAGCGCC

1801 GGAAGTGTGCAAGGGAGGCACTCCGGAGGTCCCGGTGCCCGTCCAGGGAGCAATGCGT
CCTTCACAACGTCCCTCCGTGAGGCCCTCCAGGGCGCACGGGCAGGTCCCTCGTTACGCA

1861 CCTCGGGTTCGTCCCCAGCCGCGTCTACGCGCCTCCGTCCCTCCCGTCCAGGTCCGGCATT
GGAGCCAAGCAGGGTCCGGCGCAGATGCGCGGAGGCAGGAGGGAAAGTGCAGGCCGTAA

1921 CGTGGTCCCCGGAGCCGACGCCCGTCCGGACCTGGAGGCAGCCCTGGGTCTCCGGA
GCACCAACGGGCCTCGGGCTGCGGGCGCAGGCCTGGACCTCCGTGGGACCCAGAGGCCT

1981 TCAGGCCAGCGGCCAAAGGGTCCGGCACGCACCTGTTCCAGGGCCTCCACATCATGGC
AGTCGGTCCGGTTTCCAGCGGGTGTGGACAAGGGTCCGGAGGTGTAGTACCG

2041 CCCTCCCTCGGGTTACCCACAGCCTAGGCCATTGACCTCTCTCCGCTGGGGCCCTCG
GGGAGGGAGCCAATGGGTGTCGGATCCGGCTAACGCTGGAGAGAGGCACCCGGGAGC

Sp1

2101 CTGGCGTCCCTGCACCCCTGGGAGCGCGAGCGGGCGCGCGGGGGAAAGCGCGGCCAGAC
GACCGCAGGGACGTGGACCCCTCGCGCTGCCCGCGCCCCCTCGCGCCGGTCTG

2161 CCCCCGGTCCGCCGGAGCAGCTGCCTGTCGGGGCCAGGCCGGCTCCAGTGGATTG
GGGGCCCAGGCGGGCTCGTCGACCGACAGCCCGGTCCGGCCGGAGGGTACCTAACG

2221 CGGGCAACAGACGCCAGGACCGCGCTTCCCACGTGGGGAGGGACTGGGGACCCGGGCA
GCCCGTTGTCTGCGGGCTCTGGCGCGAACGGGTGCACCGCCTCCCTGACCCCTGGGGCGT

Sp1

=====

E2F

2281 CGGGTCCCTGCCCTTCACCTCCAGCTCCGCCCTCGTCCGGCGGAACCCCGCCCCGTCCC
GGCCAGGACGGGGAAAGTGGAAAGGTGAGGCGAGCAGGCCGCGCTGGGGCGGGCAGGG

2341 GAACCCTTCCGGGTCCCCGGCCAGCCCTTCCGGGCCATCCCAGCCGTCCCGTTCT
CTTGGGAAGGGCCCAGGGCGGGTGGGGAGGGCCGGTAGGGTCGGGCAGGGCAAGGA

Sp1

=====

E2F

NFkB

2401 TTTCCGCGCCCCCGCCCTCTCTCGEGGGEGCGAGTTTCAGGCAGCGCTCGCTCTGCTG
AAAGGCGCCGGGGCGGGAGAGGAGCGCCGCGCTAACAGTCCGTCGCGACGCAGGACGACG

hTRT5'

*****>

2461 GCACGTGGGAAGCCCTGGCCCCGGCACCCCGCGATGCCGCGCTCCCGCTGCCGAG
CGTGCACCCCTCGGGACCGGGCGCTACGGCGCGAGGGCGACGGCT

2521 CCGTGCCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTGTG
GGCACCGAGGGACGACGCGTCGGTGTGGCGCTCACGACGGCACCAGGTGCAAGCAG

FIG. 21
(CONTINUED)

+

+

31/103

E2F

2581 GGCGCCTGGGGCCCAGGGCTGGCGCTGGTGCAGCGGGGACCCGGCGCTTCGCG
CCCGGGACCCGGGTCCCGACCGCGACCACGTGCGCCCCCTGGGCCGAAAGGCGC

*

2641 CGCTGGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACCCACGGCCGGCCCCCGCG
GCGACCACCGGGTACGGACCACACGCACGGGACCTGCGTGCCGGGGCGCGGG

NFkB

=====

2701 CCTCCTTCCGCCAGGTGGCCTCCCCGGGTGGCGTCCGGCTGGGTTGAGGGCGGCC
GGAGGAAGGCAGGTCCACCCGGAGGGGCCAGCCGAGGCCAACCTCCCAGGCC

Topo_II_cleavag

::: :

NFkB

++++++

NFkB

=====

Intron1

*****>

2761 GGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCAGCTCAGGGCGCTTCCCCCAGGT
CCCCCTTGGTCGCTGTACGCCTCTCGCTCCGCTGAGTCCCAGAAGGGGGCGTCCA

e_site

:::

2821 GTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGCGCGAA
CAGGACGGACTTCCTCGACCACGGGCTCACGACGTCTCGACACGCTCGCGCCGCGTT

2881 GAACGTGCTGGCTTCGGCTTCGCGCTGCTGGACGGGGCCGCGGGGGCCCCCGAGGC
CTTGACGACCGGAAGCCGAAGCGCGACGACCTGCCCCGGCGCCCCGGGGCTCCG

2941 CTTCACCAACCAGCGTGCAGCTACCTGCCAACACGGTGACCGACGCAGTGCAGGGGGAG
GAAGTGGTGGTCGACCGCTCGATGGACGGGTTGTGCCACTGGCTCGTGACGCCCTC

3001 CGGGCGTGGGGCTGCTGCTGCCCGCGTGGCGACGAGTGCTGGTTCACCTGCTGGC
GCCCGCACCCCGACGACGCGCGACCCGCTGCTGACGACCAAGTGGACGACCG

3061 ACGCTGCGCGCTCTTGCTGGTGGCTCCAGCTGCCCTACCAAGGTGTGCGGGCGCG
TGCGACGCGAGAAACACGACCAACCGAGGGTCACGCGGATGGTCCACACGCCCGCG

3121 GCTGTACCAAGCTCGCGCTGCCACTCAGGCCCGCCACACGCTAGTGGACCCCG
CGACATGGTCGAGCCCGACGGTGAGTCCGGGGGGCGGTGTGCGATCACCTGGGC

3181 AAGGCGTCTGGATGCGAACGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCT
TTCCCGACACCTACGCTGCCGGACCTGGTATCGCAGTCCCTCCGGCCCCAGGGGA

3241 GGGCCTGCCAGCCCCGGGTGCGAGGAGGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTT
CCCGGACGGTCGGGGCCACGCTCCGCGCCCCGTACGGTCGGCTTCAGACGGCAA

3301 GCCCAAGAGGCCAGGCAGTGGCGCTGCCCTGAGCCGGAGCGGACGCCGTTGGGAGGG
CGGGTTCTCGGGTCCCGCACCGCGACGGGACTCGGCCTCGCCTGCCGAAACCGTCCC

FIG. 21
(CONTINUED)

+

+

32/103

3361 GTCCTGGGCCACCGGGCAGGACGCGTGGACCGAGTGACCGTGGTTCTGTGTGGTGT
CAGGACCCGGTGGGCCCTGCACCTGGCTACTGGCACAAAGACACACCACAG

3421 ACCTGCCAGACCCGCCGAAGAAGCCACCTCTTGAGGGTGCCTCTGGCACGCGCA
TGGACGGTCTGGCGGCTCTCGGTGGAGAACCTCCCACGCGAGAGACCGTGCACGGT

3481 CTCCCACCCATCCGTGGGCCAGCACACGCCGGCCCCCATCCACATCGCGGCCACC
GAGGGTGGTAGGCACCCGGCGTGTGGTGCACGGGGGTAGGTGTAGCGCCGGTGG

3541 ACGTCCCTGGGACACGCCCTGTCCCCCGGTGTACGCCAGACCAAGCAGTCTCTACTC
TGCAGGGACCCCTGTGCGAACAGGGGCCACATGCGCTCTGGTCGTGAAGGAGATGAG

3601 CTCAGGCACAAGGAGCAGCTGCGGCCCTCCTCCTACTCAGCTCTGAGGGCCAGCCT
GAGTCGCTGTCCTCGTACGCCGGAGGAAGGATGAGTCGAGAGACTCCGGTCGGA

3661 GACTGGCGCTCGGAGGCTGTGGAGACCATCTTCTGGGTCCAGGCCCTGGATGCCAGG
CTGACCGCAGCCTCCGAGCACCTCGTAGAAAAGACCAAGGTCCGGACCTACGGTCC

3721 GACTCCCCGAGGTGCCCCGCCTGCCAGCGTACTGGCAAATGCCGCCCTGTTCT
CTGAGGGCGTCCAACGGGGCGGACGGGTCGCGATGACCGTTACGCCGGGACAAAGA

3781 GGAGCTGCTTGGAACCAACGCGCAGTGCCCCTACGGGTGCTCCTCAAGACGCACTGCC
CTCGACGAACCCCTGGTGCACGGGATGCCACAGACACGGGCCCTTCGGGAGCTGG

3841 GCTGCAGCTGCGGTACCCAGCAGCCGGTGTCTGTGCCCGGGAGAAGCCCAGGGCTC
CGACGCTCGACGCCAGTGGGTCGTCGGCCACAGACACGGGCCCTTCGGGAGCTGG

3901 TGTGGCGGCCCGAGGAGGAGGACACAGACCCCGTCGCTGGTGCAGCTGCTCCGCCA
ACACCGCCGGGGCTCCTCCTGTGTCTGGGGCAGCGGACCGTCGACGAGGGGGT

3961 GCACAGCAGCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCCTGCGCCGGCTGGTGGC
CGTGTGTCGGGACCGTCCACATGCCGAAGCACGCCGGACGGACGCCGGACACGG

4021 CCCAGGCCTCTGGGCTCCAGGCACAACGAACGCCGCTCCCTCAGGAACACCAAGAAGTT
GGTCGGAGACCCGAGGTCCGTGTTGCTGCGCGAAGGAGTCCTGTGGTTCTCAA

4081 CATCTCCCTGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGT
GTAGAGGGACCCCTCGTACGGTTCGAGAGCGACGTCCACTGCACCTTCTACTCGCA

4141 GCGGGACTGCGCTGGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTGAGGGGCC
CGCCCTGACCGAACCGACGCGTCCTCGGGTCCACTCCTCCACCAACCGGAGCTCCGGG

Intron2

4201 AGGCCCTAGAGCTGAATGAGTAGGGGCTCAGAAAAGGGGGCAGGAGGCCCTGGTCT
TCCGGGTCTCGACTTACGTCATCCCCGAGTCTTCCCCGTCCGTCTCGGGACAGAGGA

4261 CCTGTCTCCATCGTCACGTGGGCACACGTGGTTTCGTCAGGACGTCGAGTGGACACG
GGACAGAGGTAGCAGTGCACCGTGTGCAACGAAAAGCGAGTCCTGCAGCTCACCTGTGC

4321 *****>
GTGATCGAGGTGAC
CACTAGCTCCAGCTG

FIG. 21
(CONTINUED)

+

+

33/103

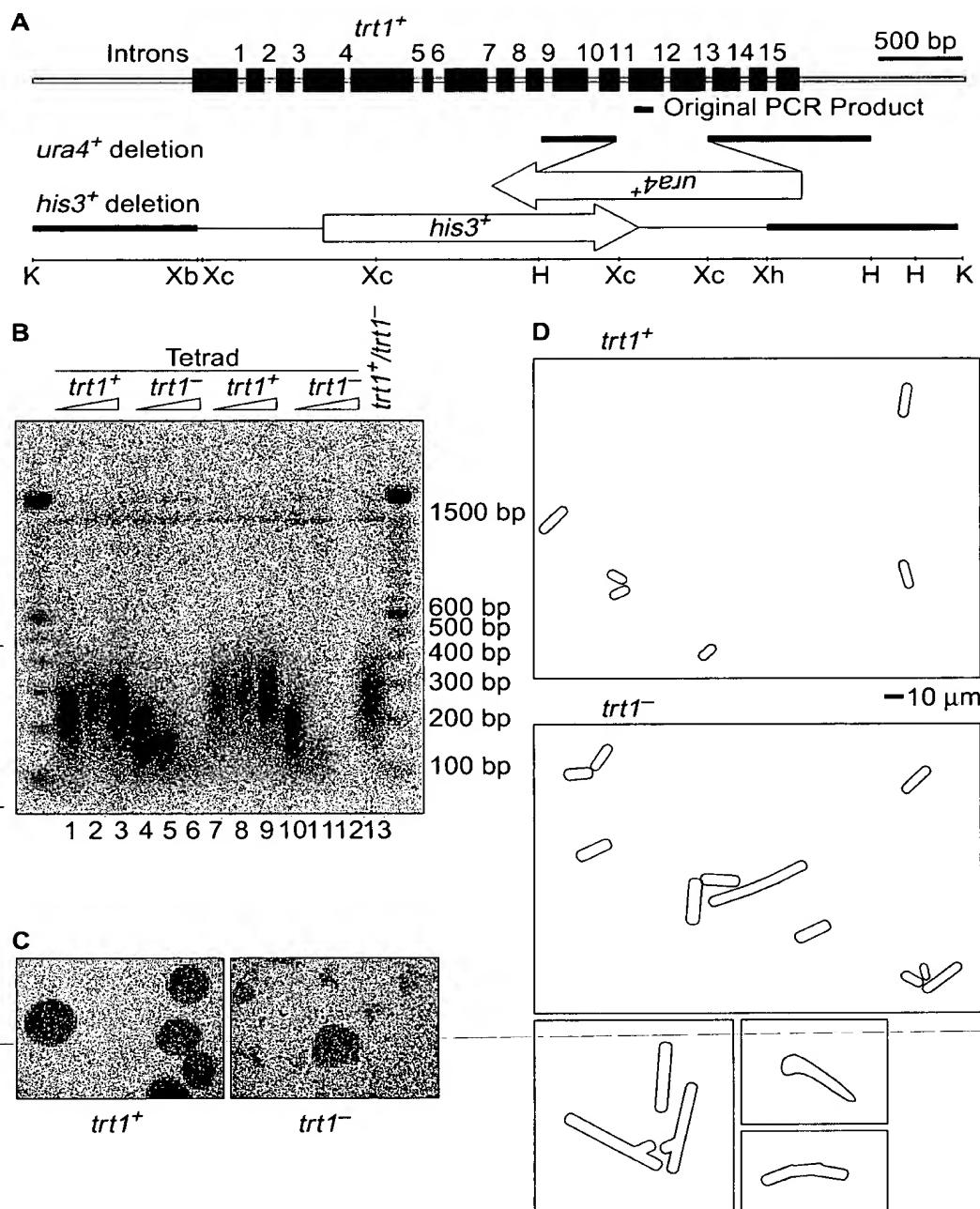


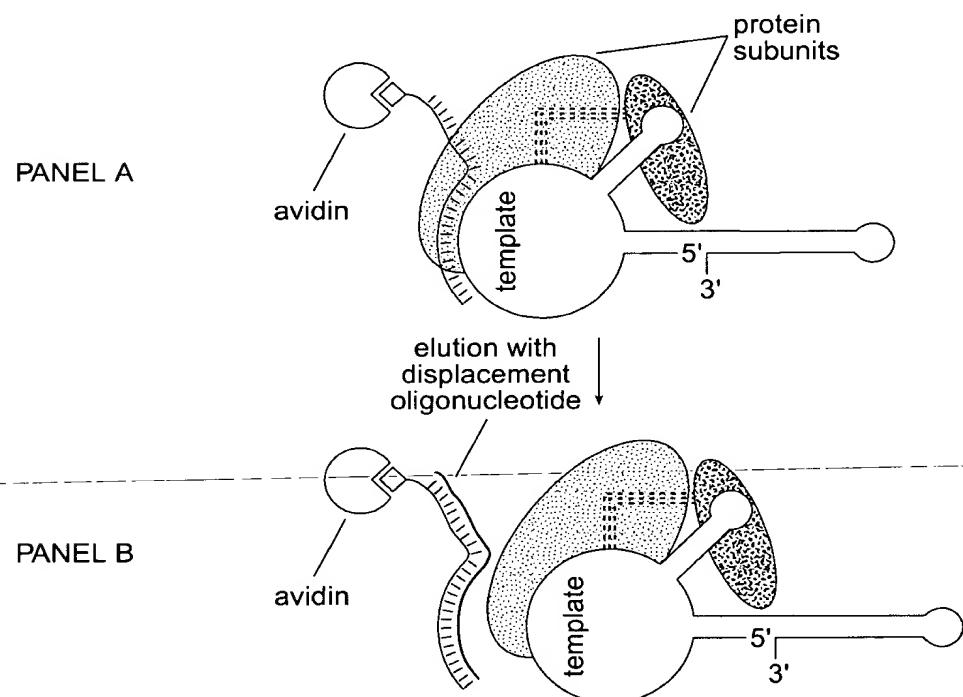
FIG. 22

+

gccaagttcctgcactggctgatgagtgtacgtcgctcgagctgctcaggctttctt
 tatgtcacggagaccacgtttcaaaagaacaggcttttctaccggaaagagtgtctgg
 agcaagttgcaaaggcattggaatcagacacgcacttgaagaggggtgcagctgcgggacgtg
 tcggaaggcagggcaggcaggcatcgggaaagccaggcccgcctgctgacgtccagactc
 cgcttcatccccaaaggcctgacgggctgcggccgattgtgaacatggactacgtcgtggga
 gccagaacgttcccgagaaaaagagggcgcagcgtcacctcgagggtgaaggcactg
 ttcagcgtgctcaactacgagcgggcgcg

FIG. 23

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTACCTGCAGGAG
 ACCAGCCCGCTGAGGGATGCCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCC
 AGCAGTGGCCTTCGACGTCTCCTACGCTTACATGTGCCACCACGCCGTGCGCATC
 AGGGGCAAGTC

FIG. 24**FIG. 26**

+

35/103

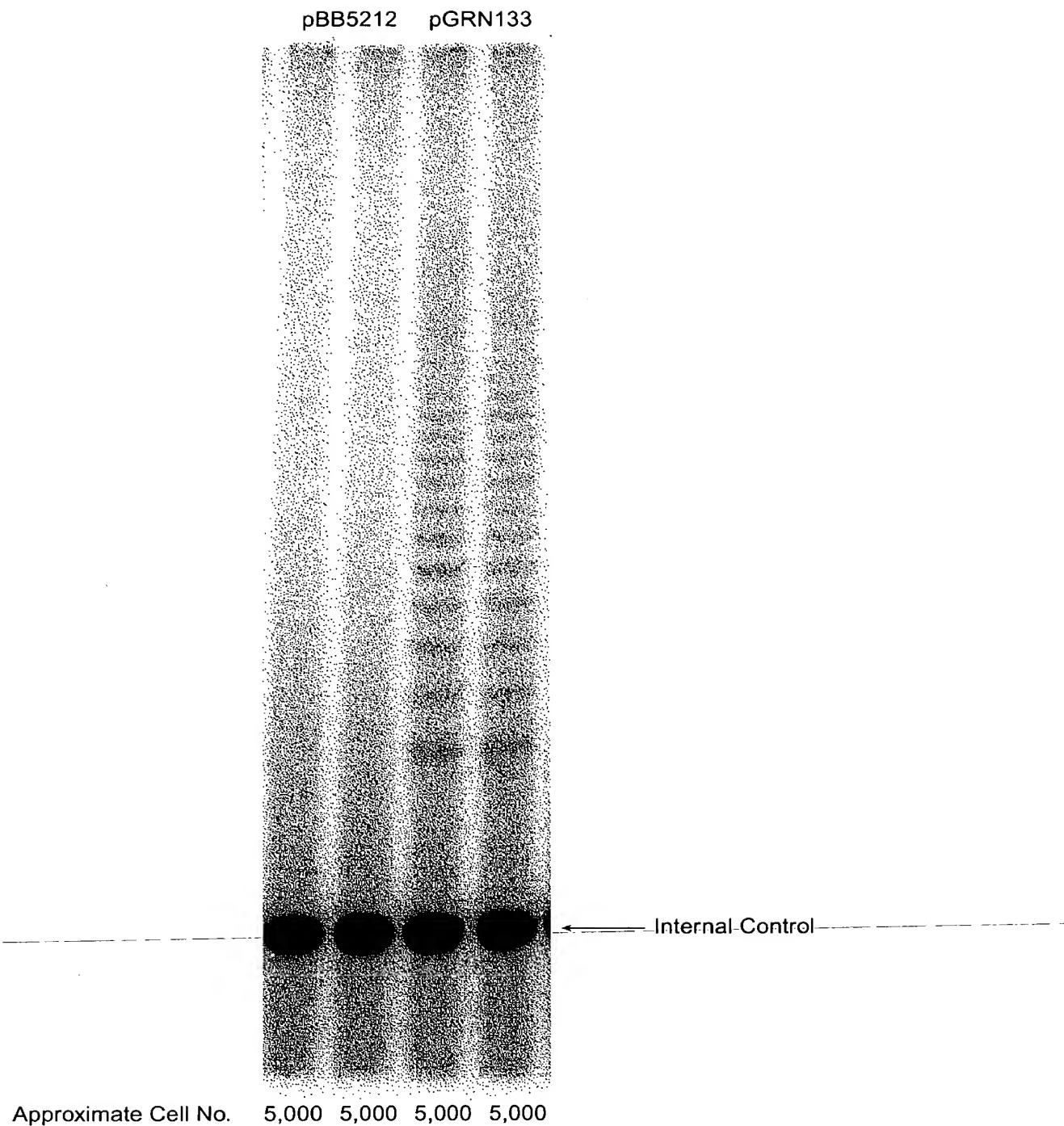
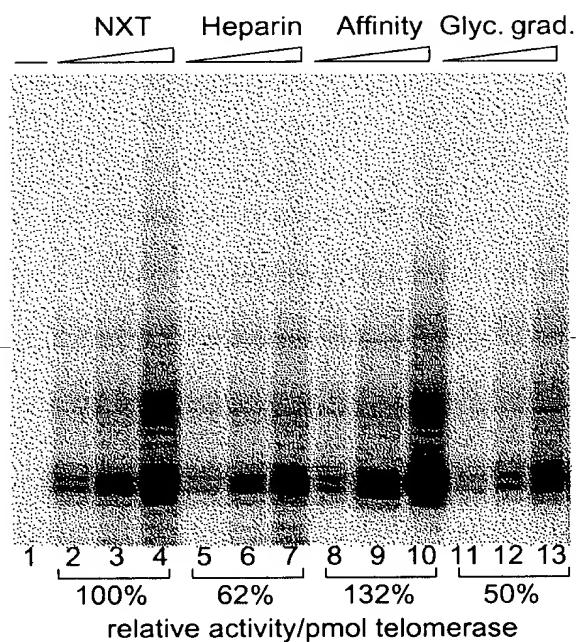
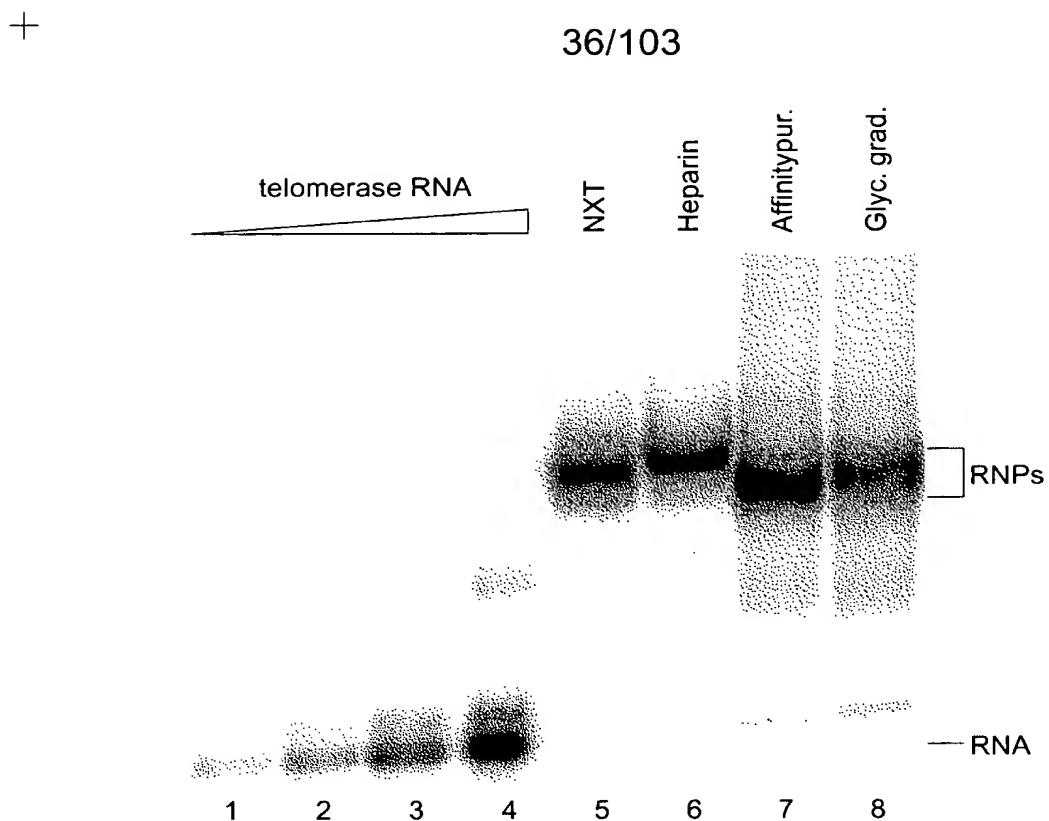


FIG. 25

+



+

37/103

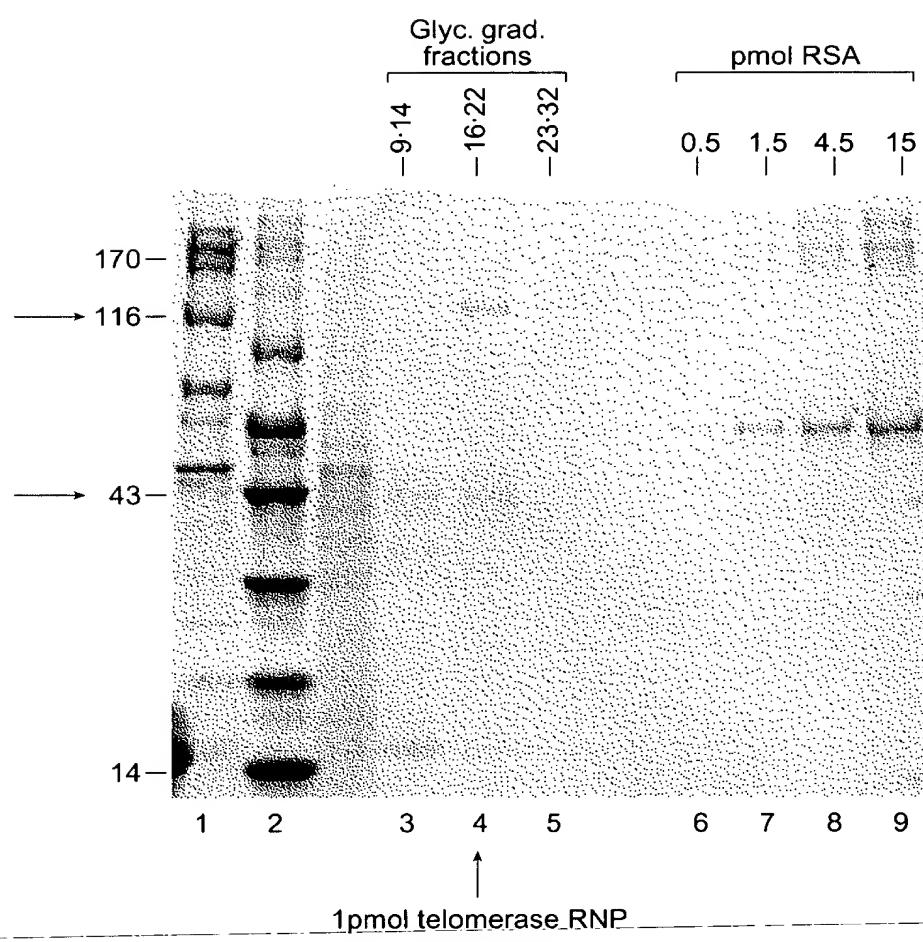


FIG. 29

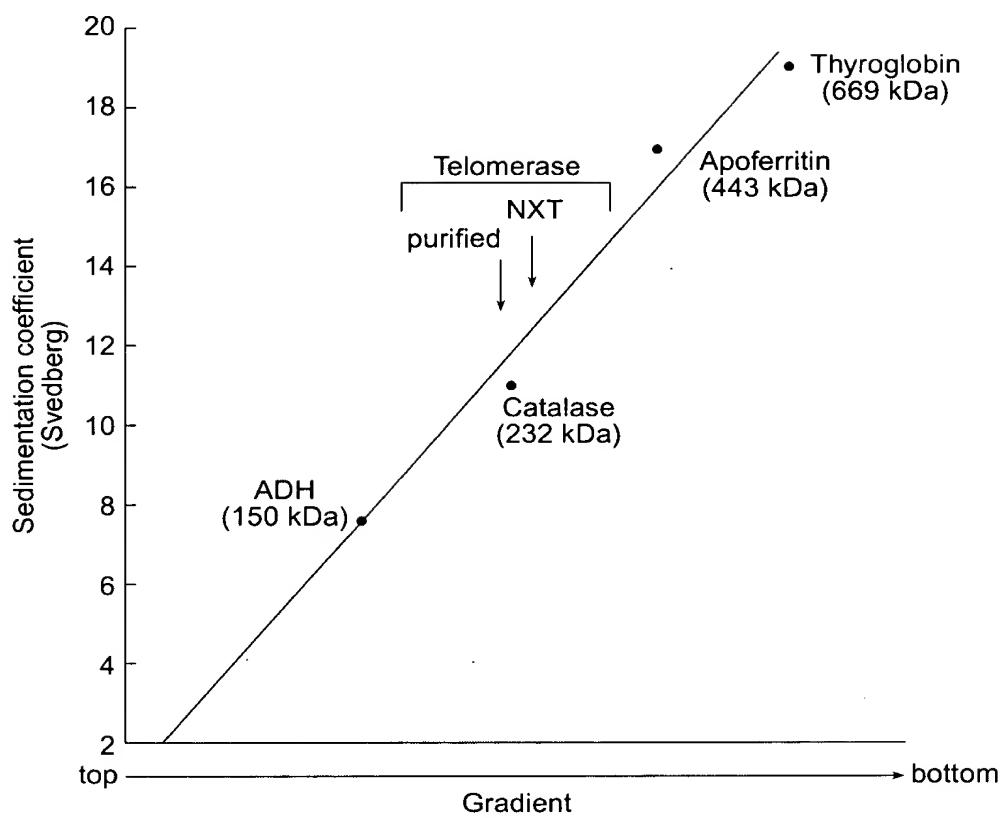


FIG. 30

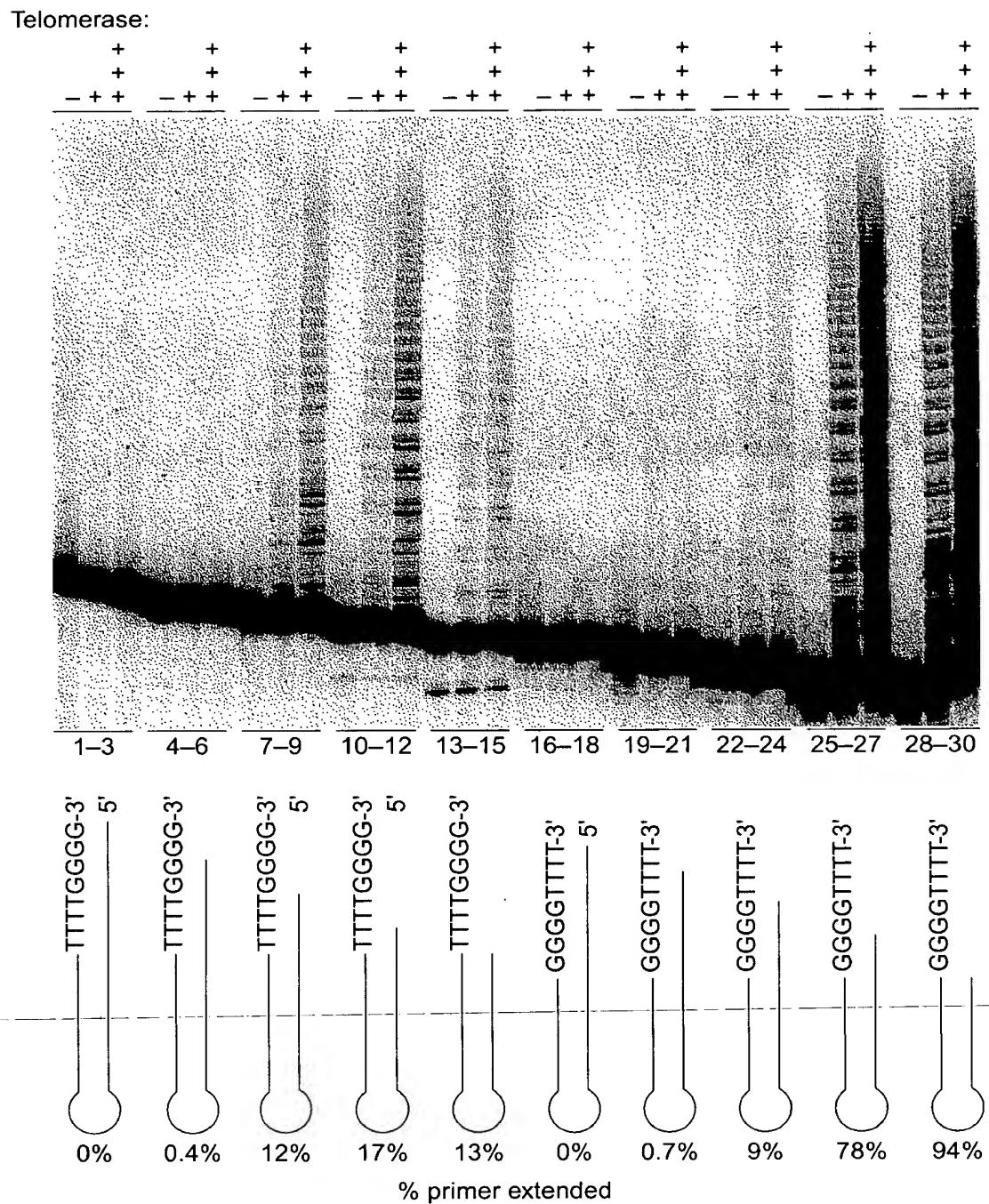


FIG. 31

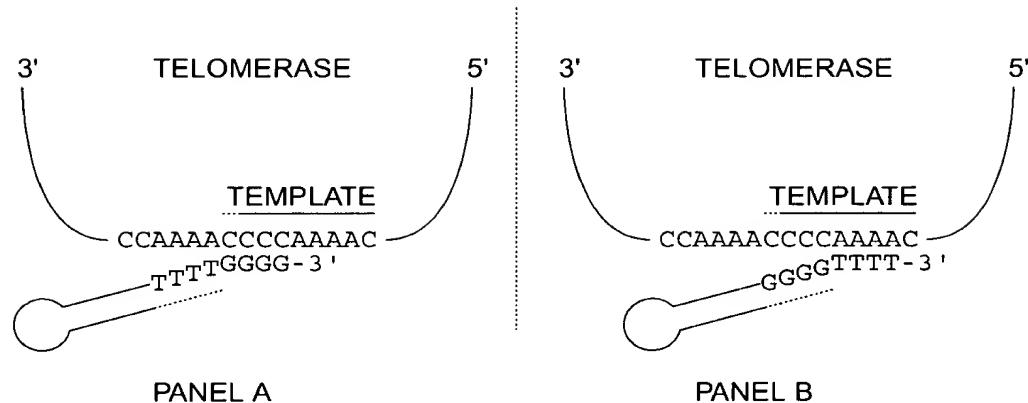


FIG. 32

1	CCCCAAAACC	CCAAAACCCC	AAAACCCCTA	TAAAAAAAGA	AAAAATTGAG
51	GTAGTTAGA	AATAAAATAT	TATTCCCGCA	CAAATGGAGA	TGGATATTGA
101	TTGGGATGAT	ATAGAAAATT	TACTTCCTAA	TACATTCAAC	AAGTATAGCA
151	GCTCTTGAG	TGACAAGAAA	GGATGCAAAA	CATTGAAATC	TGGCTCGAAA
201	TCGCCTTCAT	TGACTATTCC	AAAGTTGCAA	AAACAATTAG	AGTTCTACTT
251	CTCGGATGCA	AATCTTATA	ACGATTCTTT	CTTGAGAAAA	TTAGTTTAA
301	AAAGCGGAGA	GCAAAGAGTA	GAAATTGAAA	CATTACTAAT	GTAAATAATAA
351	AATCAGGTAA	TGAGGATTAT	TCTATTTTTT	AGATCACTTC	TTAAGGAGCA
401	TTATGGAGAA	AATTACTTAA	TACTAAAAGG	TAACACAGTTT	GGATTATTTTC
451	CCTAGCCAAC	AATGATGAGT	ATATTAAATT	CATATGAGAA	TGAGTCAAAG
501	GATCTCGATA	CATCAGACIT	ACCAAAGACA	AACTCGCTAT	AAAACGCAAG
551	AAAAAGTTTG	ATAATCGAAC	AGCAGAAAGAA	CTTATTGCAT	TTACTATTGCG
601	TATGGGTTTT	ATTACAATTG	TTTTAGGTAT	CGACGGTGAA	CTCCCGAGTC
651	TTGAGACAAAT	TGAAAAAGCT	GTTTACAACAT	GAAGGAATCG	CAGTTCTGAA
701	AGTTCTGATG	TGTATGCCAT	TATTTTGTA	ATTAATCTCA	AATATCTTAT
751	CTCAATTAA	TGGATAGCTA	TAGAAACAAA	CCAAATAAAC	CATGCAAGTT
801	TAATGGAATA	TACGTTAAAT	CCTTTGGGAC	AAATGCACAC	TGAATTTATA
851	TTGGATTCTT	AAAGCATAGA	TACACAGAAT	GCTTTAGAGA	CTGATTTAGC
901	TTACAACAGA	TTACCTGTTT	TGATTACTCT	TGCTCATCTC	TTATATCTTT
951	AAAAGAAGCA	GGGAAATGA	AAAGAAGACT	AAAGAAAGAG	ATTTCAAAAT
1001	TTGTTGATTC	TTCTGTAACC	CCAATTAAACA	ACAAGAATAT	TAGCAACGAA
1051	AAAGAAGAAG	AGCTATCACA	ATCCTGATTC	TTAAAGATT	AAAAAATTCC
1101	AGGTAAAGAGA	GATACATTCA	TTAAAATTCA	TATATTATAG	TTTTTCATTT
1151	CACAGCTGTT	ATTTCTTTT	ATCTTAACAA	TATTTTTGTA	TTAGCTGGAA
1201	GTAAAAAGTA	TCAAAATAGA	GAAGCGCTAG	ACTGAGGTAA	CTTAGCTTAT
1251	TCACATTCCAT	AGATCGACCT	TCATATATCC	AATACGATGA	TAAGGAAACA
1301	GCAGTCATCC	GTTTTAAAAA	TAGTGCTATG	AGGACTAAAT	TTTTAGAGTC
1351	AAGAAATGGA	GCCGAAATCT	TAATCAAAA	GAATTGCGTC	GATATTGCAA
1401	AAGAATCGAA	CTCTAAATCT	TTCGTTAATA	AGTATTACCA	ATCTTGATTG
1451	ATTGAAGAGA	TTGACGAGGC	AACTGCACAG	AAGATCATT	AAGAAATAAA
1501	GTAACCTTTA	TTAATTAGAG	AATAAACTAA	ATTACTAATA	TAGAGATCAG
1551	CGATCTCAA	TTGACGAAAT	AAAAGCTGAA	CTAAAGTTAG	ACAATAAAAA
1601	ATACAAACCT	TGGTCAAAAT	ATTGAGGAAG	GAAAAGAAGA	CCAGTTAGCA
1651	AAAGAAAAAA	TAAGGCAATA	AATAAAATGA	GTACAGAAAGT	GAAGAAATAA
1701	AAGATTATTA	TTTTCAATA	ATTATTGAA	AAGAGGGTT	TTGGGGTTTT
1751	GGGGTTTGG	GG			

FIG. 34

+

41/103

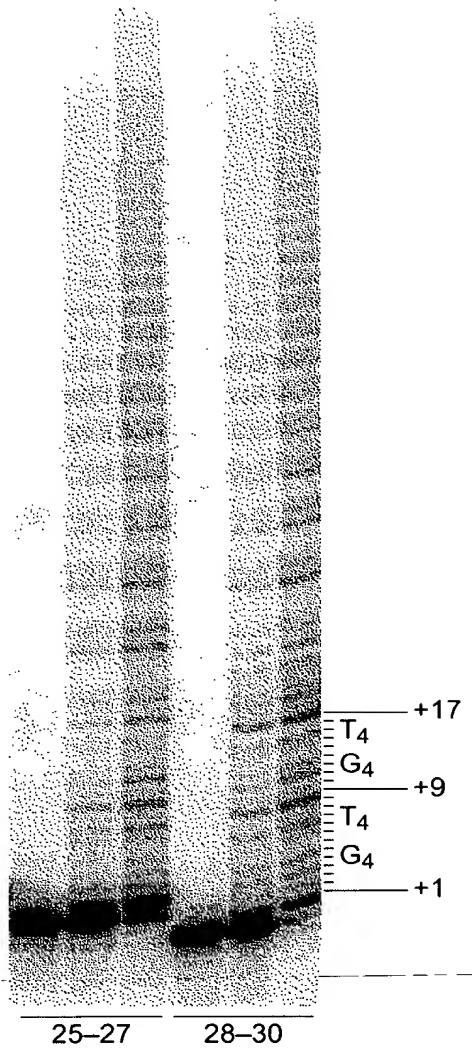


FIG. 33

+

1 CCCCCAAAACCCAAAAACCCAAAACCCCTATAAAGAAAAATTGAGGTAGTTAGA 60
 1 GGGGTTTGGGGTTTGGGATTTCTTTTAACTCATCAAATCT
 a P Q N P K T P K P L * K K K K L R * F R -
 b P K T P K P Q N P Y K K R K N * G S L E -
 c P K P Q N P K T P I K K E K I E V V * K -
 AATAAAATATTATTCCCACAAATGGAGATGGATATTGATTTGGATGATAGAAAATT
 61 TTATTTATAATAAGGGCGTGTACCTACCTATAACTAACCTACTATATCTTTAA 120
 a N K I L F P H K W R W I L I W M I * K I -
 b I K Y Y S R T N G D G Y * F G * Y R K F -
 c * N I I P A Q M E M D I D L D D I E N L -
 TACTTCCTAATACATTCAACAAGTATAGCAGCTCTGTAGTGACAAGAAAGGATGCAAAA
 121 ATGAAGGATTATGTAAGTTGTCATATCGTCGAGAACATCACTGTTCTTCCTACGTTT 180
 a Y F L I H S T S I A A L V V T R K D A K -
 b T S * Y I Q Q V * Q L L * C Q E R M Q N -
 c L P N T F N K Y S S S C S D K K G C K T -
 CATTGAAATCTGGCTCGAAATGCCCTTCATTGACTATTCCAAAGTTGCAAAACAATTAG
 181 GTAACTTTAGACCGAGCTTAGCGGAAGTACTGATAAGGTTCAACGTTTTGTTAAC 240
 a H * N L A R N R L H * L F Q S C K N N * -
 b I E I W L E I A F I D Y S K V A K T I R -
 c L K S G S K S P S L T I P K L Q K Q L E -
 AGTTCTACTTCTCGGATGCAAATCTTATAACGATTCTTCTTGAGAAAATTAGTTAA
 241 TCAAGATGAAGAGCCTACGTTAGAAATATTGCTAAGAAAGAACTCTTTAACAAAATT 300
 a S S T S R M Q I F I T I L S * E N * F * -
 b V L L L G C K S L * R F F L E K I S F K -
 c F Y F S D A N L Y N D S F L R K L V L K -
 AAAGCGGAGAGCAAAGAGTAGAAATTGAAACATTACIATGTTAAATAAAATCAGGTAA
 301 TTTCGCCTCTCGTTCTCATCTTAACTTGTAATGATTACAAATTATTTAGTCCATT 360
 a K A E S K E * K L K H Y * C L N K I R * -
 b K R R A K S R N * N I T N V * I K S G N -
 c S G E Q R V E I E T L L M F K * N Q V M -
 TGAGGATTATTCTATTAGATCACTCTTAAGGAGCATTATGGAGAAAATTACTTAA
 361 ACTCCTAATAAGATAAAAATCTAGTGAAGAATTCTCGTAATACCTCTTTAACATT 420
 a * G L F Y F L D H F L R S I M E K I T * -
 b E D Y S I F * I T S * G A L W R K L L N -
 c R I I L F F R S L L K E H Y G E N Y L I -

FIG. 35

43/103

TACTAAAAGGTAAACAGTTGGATTATTCCTAGCCAACAATGATGAGTATATTAAATT
 421 ATGATTTCCATTTGCAACCTAATAAAGGGATCGGTTACTACTCATATAATTAA 480

a Y * K V N S L D Y F P S Q Q * * V Y * I -
 b T K R * T V W I I S L A N N D E Y I K F -
 c L K G K Q F G L F P * P T M M S I L N S -

CATATGAGAATGAGTCAAAGGATCTGATACATCAGACTTACCAAAGACAAACTCGCTAT
 481 GTATACTCTTACTCAGTTCTAGAGCTATGAGTCTGAATGGTTCTGTTGAGCGATA 540

a H M R M S Q R I S I H Q T Y Q R Q T R Y -
 b I * E * V K G S R Y I R L T K D K L A I -
 c Y E N E S K D L D T S D L P K T N S L *

AAAACGCAAGAAAAAGTTGATAATCGAACAGCAGAAGAACTTATTGCATTTACTATCG
 541 TTTTGCGTCTTTCAAACATTAGCTTGCGTCTTGAATAACGTAATGATAAGC 600

a K T Q E K V * * S N S R R T Y C I Y Y S -
 b K R K K K F D N R T A E E L I A F T I R -
 c N A R K S L I I E Q Q K N L L H L L F V -

TATGGGTTTATTACAATTGTTTAGGTATCGACGGTAACTCCGAGTCTGAGACAAT
 601 ATACCCAAAATAATGTTAACAAATCCATAGCTGCCACTGAGGGCTCAGAACTCTGTTA 660

a Y G F Y Y N C F R Y R R * T P E S * D N -
 b M G F I T I V L G I D G E L P S L E T I -
 c W V L L Q L F * V S T V N S R V L R Q L -

TGAAAAAGCTGTTACAACGTGAAGGAATCGCAGTTCTGAAAGTTCTGATGTGTATGCCAT
 661 ACTTTTGCACAAATGTTGACTTCCTTAGCGTCAAGACTTCAAGACTACACATACGGTA 720

a * K S C L Q L K E S Q F * K F C C V C H -
 b E K A V Y N * R N R S S E S S D V Y A I -
 c K K L F T T E G I A V L K V L M C M P L -

TATTTGTGAATTAATCTCAAATATCTTATCTCAATTAAATGGATAGCTATAGAAACAAA
 721 ATAAAACACTTAATTAGAGTTATAGAATAGAGTTAAATTACCTATCGATATCTTGT 780

a Y F V N * S Q I S Y L N L M D S Y R N K -
 b I L * I N L K Y L I S I * W I A I E T N -
 c F C E L I S N I L S Q F N G * L * K Q T -

CCAAATAAACCATGCAAGTTAATGGAATATACGTTAAATCCTTGGGACAAATGCACAC
 781 GGTTTATTGGTACGTTCAAATTACCTTATGCAATTAGGAAACCTGTTACGTGTG 840

a P N K P C K F N G I Y V K S F G T N A H -
 b Q I N H A S L M E Y T L N P L G Q M H T -
 c K * T M Q V * W N I R * I L W D K C T L -

TGAATTATATTGGATTCTAAAGCATAGATACACAGAATGCTTAGAGACTGATTAGC
 841 ACTTAAATATAACCTAAGAACCTGTTACGAAATCTCTGACTAAATCG 900

a * I Y I G F L K H R Y T E C F R D * F S -
 b E F I L D S * S I D T Q N A L E T D L A -
 c N L Y W I L K A * I H R M L * R L I * L -

FIG. 35
(CONTINUED)

44/103

TTACAACAGATTACCTGTTTGATTACTCTGCTCATCTCTTATATCTTAAAGAAGCA
 901 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 960
 AATGTTGTCTAATGGACAAAACATAATGAGAACGAGTAGAGAATATAGAAATTTCTTCGT

a L Q Q I T C F D Y S C S S L I S L K E A -
 b Y N R L P V L I T L A H L L Y L * K K Q -
 c T T D Y L F * L L L L I S Y I F K R S R -

GGCGAAATGAAAAGAAGACTAAAGAAAGAGATTCAAAATTGTTGATTCTTCTGTAACC
 961 -----+-----+-----+-----+-----+-----+-----+-----+ 1020
 CCGCTTTACTTTCTTCTGATTCTTCTAAAGTTAAACAACTAAGAACATTGG

a G E M K R R L K K E I S K F V D S S V T -
 b A K * K E D * R K R F Q N L L I L L * P -
 c R N E K K T K E R D F K I C * F F C N R -

GGAATTAACAACAAGAATATTAGCAACGAAAAAGAAGAAGAGCTATCACAACTGATT
 1021 -----+-----+-----+-----+-----+-----+-----+-----+ 1080
 CCTTAATTGTTGTTCTTATAATCGTTGCTTTCTTCTCGATAGTGTAGGACTAAG

a G I N N K N I S N E K E E E L S Q S * F -
 b E L T T R I L A T K K K K S Y H N P D S -
 c N * Q Q E Y * Q R K R R R A I T I L I L -

TTAAAGATTCAAAATTCCAGGTAAGAGAGATACATTCAATTAAATTCATATATTAG
 1081 -----+-----+-----+-----+-----+-----+-----+-----+ 1140
 AATTCTAAAGTTTAAGGTCCATTCTCTATGTAAGTAATTAAAGTATATAATATC

a L K I S K I P G K R D T F I K I H I L * -
 b * R F Q K F Q V R E I H S L K F I Y Y S -
 c K D F K N S R * E R Y I H * N S Y I I V -

TTTTTCATTCACAGCTGTTATTTCTTATCTTAACAATATTTTGATTAGCTGGAA
 1141 -----+-----+-----+-----+-----+-----+-----+-----+ 1200
 AAAAGTAAAGTGTGACAAATAAGAAAAGAGATAGAATTGTTATAAAACTAATCGACCTT

a F F I S Q L L F S F I L T I F F D * L E -
 b F S F H S C Y F L L S * Q Y F L I S W K -
 c F H F T A V I F F Y L N N I F * L A G S -

GTAAAAAGTATCAAATAAGAGAAGCGCTAGACTGAGGTAACTTAGCTTATTCACATTCA
 1201 -----+-----+-----+-----+-----+-----+-----+-----+ 1260
 CATTTTCATAGTTATTCTCTCGCGACTCCATTGAATCGAATAAGTGTAGTA

a V K S I K * E K R * T E V T * L I H I H -
 b * K V S N K R S A R L R * L S L F T F I -
 c K K Y Q I R E A L D * G N L A Y S H S * -

AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTAAAAA
 1261 -----+-----+-----+-----+-----+-----+-----+-----+ 1320
 TCTAGCTGGAAGTATAGGTATGCTACTATTCCCTTGTGTCAGTAGGCAGAAATT
 TTT

a R S T F I Y P I R * * G N S S H P F * K -
 b D R P S Y I Q Y D D K E T A V I R F K N -
 c I D L H I S N T M I R K Q Q S S V L K I -

TAGTGCTATGAGGACTAAATTAGAGTCAGAAATGGAGCCGAAATCTTAATCAAAA
 1321 -----+-----+-----+-----+-----+-----+-----+-----+ 1380
 ATCACGATACTCCTGATTAAAACTCAGTTACCTCGGCTTAGAATTAGTTT

a * C Y E D * I F R V K K W S R N L N Q K -
 b S A M R T K F L E S R N G A E I L I K K -
 c V L * G L N F * S Q E M E P K S * S K R -

FIG. 35
 (CONTINUED)

1381 GAATTGCCTCGATATTGCAAAAGAATCGAACTCTAAATCTTCGTTAATAAGTATTACCA 1440
 CTTAACGCAGCTATAACGTTCTTAGCTGAGATTAGAAAGCAATTATTCATAATGGT

a	E	L	R	R	Y	C	K	R	I	E	L	*	I	F	R	*	*	V	L	P	-
b	N	C	V	D	I	A	K	E	S	N	S	K	S	F	V	N	K	Y	Y	Q	-
c	I	A	S	I	L	Q	K	N	R	T	L	N	L	S	L	I	S	I	T	N	-

ATCTTGATTGATTGAAGAGATTGACGAGGCCACTGCACAGAAGATCATTAAAGAAATAAA 1441 1500
 1441 TAGAACTAACTAACTCTCTAACTGCTCCGTTGACGTGCTTAGTAATTCTTTATTT

a	I	L	I	D	C	R	D	C	R	G	N	C	T	E	D	H	*	R	N	K	-
b	S	*	L	I	E	E	I	D	E	A	T	A	Q	K	I	I	K	E	I	K	-
c	L	D	*	L	K	R	L	T	R	Q	L	H	R	R	S	L	K	K	*	S	-

GTAACTTTTATTAATTAGAGAATAAACTAAATTACTAATATAGAGATCAGCGATCTCAA 1501 1560
 1501 CATTGAAAATAATTAATCTCTTATTTGATTAAATGATTATCTCTAGTCGCTAGAAGTT

a	V	T	F	I	N	*	R	I	N	*	I	T	N	I	E	I	S	D	L	Q	-
b	*	L	L	L	I	R	E	*	T	K	L	L	I	*	R	S	A	I	F	N	-
c	N	F	Y	*	L	E	N	K	L	N	Y	*	Y	R	D	Q	R	S	S	I	-

TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAACACAAACCTGGTCAAAAT 1561 1620
 1561 AACTGCTTATTTGACTTGATTCAATCTGTTATTTTATGTTGGAACCGAGTTA

a	L	T	K	*	K	L	N	*	S	*	T	I	K	N	T	N	L	G	Q	N	-
b	*	R	N	K	S	*	T	K	V	R	Q	*	K	I	Q	T	L	V	K	I	-
c	D	E	I	K	A	E	L	K	L	D	N	K	K	Y	K	P	W	S	K	Y	-

ATTGAGGAAGGAAAAGAAGACCAGTTAGCAAAAGAAAAATAAGGCAATAAAATAAGA 1621 1680
 1621 TAACTCCTCCTTTCTCTGGCAATCGTTTCTTTTATTCCGTTATTTATTTACT

a	I	E	E	G	K	E	D	Q	L	A	K	E	K	I	R	Q	*	I	K	*	-
b	L	R	K	E	K	K	T	S	*	Q	K	K	K	*	G	N	K	*	N	E	-
c	*	G	R	K	R	R	P	V	S	K	R	K	N	K	A	I	N	K	M	S	-

GTACAGAAAGTGAAGAAAATAAAAGATTATTTTTCAATAATTATTGAAAAGAGGGTT 1681 1740
 1681 CATGTCTTCACTTCTTATTTCTAAATAAAAGTTATTAAATACTTTCTCCCCAA

a	V	Q	K	*	R	N	K	R	F	I	F	F	N	N	L	L	K	R	G	V	-
b	Y	R	S	E	E	I	K	D	L	F	F	S	I	I	Y	*	K	E	G	F	-
c	T	E	V	K	K	*	K	I	Y	F	F	Q	*	F	I	E	K	R	G	F	-

TTGGGGTTTGGGGTTTGGGG
 1741 1762
 AACCCCAAAACCCCAAAACCC

a	L	G	F	W	G	F	G	-
b	W	G	F	G	V	L	G	-
c	G	V	L	G	F	W	-	

FIG. 35
 (CONTINUED)

2 EVDVQNQADNHGIHSALKTCEEIKEAKTLYSWIQKVICRCRNQSQSHYKDL 51
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 19 ELELEMQENQNDIQRVK....IDDPKQY..LVNVTAACLLQEGSYYQDK 62
 52 EDIKIFQAQTNIVATPRDYNEEDFKVIARKEVF.STGLMIELIDKCLVELL 100
 ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
 63 DERRYIITKALL....EVAESDPEFICQLAVYIRNELYIRTTNYIVAF. 107
 101 SSSDVSDRQKLQCFGQLKGNLAKTHLLTALSTQKQYFFQDEWNQVRAM 150
 .::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
 108 .CVVHKNTQPFIEKYFNKAVALLPNDLLEVCEFAQVLYI 144
 151 IGNELFRHLYTKYLIFQRTSEGTIVQFCGNNVFDHLKVNDKFDDKKQKGGA 200
 ::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
 145 FDATEFKNLY.....LDRILSQDIRKELETFRKCLQRCVRSKF 181
 201 ADMNE...PRCCSTCKYNVKNEKDHFLLNNINVPNWNNMKSRTTRIFYCHF 247
 ::|::|::|::|::|::|::|::|::|::|::|::|::|:
 182 SEFNEYQLGKYCTES..QRKKTMFRLSVTNKQWDQTKKK..... 220
 248 NRNNQFFKKHEFVSNKNNISAMDRAQTIFTNIFRFNRIRKDKVIEKI 297
 |::|::|::|::|::|::|::|::|::|::|::|::|:
 221 .RKENLLTKLQAIKESEDDKSKRETG.....DIMNVEDAIKALKPAVMKKI 264
 298 AYMЛЕKVKDFNFNYYLTKSCPLPENWRERKQKИENLINKTREEKSKYYEE 347
 |::|::|::|::|::|::|::|::|::|::|::|:
 265 AKRQNAMK.....KHMKAPKIPNSTLESKYLTTFKD 294
 348 LFSYTTDNKCVTQFINEFFYNILPKDFLTGRNRKNFQKKVKKYVELNKHE 397
 |::|::|::|::|::|::|::|::|::|::|:
 295 LIKFCHISEP.....KERVYKILGKKYPKTEEEYKAAFGDSASAPFN.PE 338
 398 LIHKNLLLEKINTREISWMQVETSAKHFYYFDHENIYVLWKLRLWIFEDL 447
 |::|::|::|::|::|::|::|::|::|::|:
 339 LAGKRMKIEISKTWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSN.. 386
 448 VVSLIRCFYVTEQQKSYSKTYYRKNIWDVIMKMSIADLKETLAEVQE 497
 |::|::|:
 387ILKAGVSD..... 394
 498 KEVEEWKKSLGFAPGKLRLIPKKTTFRPIMTFNKKIVNSDRKTTKLTTNT 547
 |::|:
 395TTHS 398
 548 KLLNSHMLKTLKNRMFKDPFGFAVFNYDDVMKKYEEFVCKWKQVGQPKL 597
 |::|:
 399 IVINK.....ICEPKAVENSKM 415
 598 FFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKNVID 647
 |::|::|::|::|::|::|::|::|:
 416 F..PLQFFSAIEAVN.EAVTKGFKAKK..RENMLKGQIEAVKE..VVE 457
 648 SKNFRKKEMKDYFRQKFQKIALEGQYPTLFSVLENEQNDLNAAKTLIVE 697
 |::|::|:
 458 KTDEEKKDM.....ELEQTEEGERVKVNEGIGKQYINSTIELAIK 496
 698 AKQRNYFKKDNLQPVINICQNYNINFNGKFYKQTKGIPQGLCVSSILSS 747
 |::|::|::|::|::|::|::|:
 497 IAVKNLDEIKGHTAIFSDVSGSMSTSMSGGAKKYGSVRTCLECALVLGL 546
 748 FYYATLEESSLGFLRDESMNPENPNVNLLMRLTDYLLITTQENNAVLFI 797
 |::|::|::|::|:
 547 MVKQRCEKSSFYIFSSPSSQCNKCYLEVDL..... 576

798	EKLINVSRENGFKFNMKK . LQTSFPLSPSKFAKYGMDSVEEQNIVQDYCD	846
577 PGDELRPSMQKLLQEKGKLGGS . TDFPYECIDEWTKNKTHVD	617
847	WIGISIDMKTALMPNINLRIEGILCTLNLNMQTKKASMWLKKKLKSFLM	896
618 NIVILSDMMIAEGYSDINVRGSSIVNSI .. KKYKDEVN	653
897	NNITHYFRKTITTEDFANKTLNKLFISSGGYKYMCAKEYKD . HFKNLAM	945
654 PNIKIF . AVDLEGYG .. KCLNLGDEFNENNYIKIFGM	687
946	SSMIDLEVSKIIYSVTRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHFIE	995
688 SDSI .. LKFISAKQGGA .. NMVE	706
996	IFSTKKYIFNRVC 1008	
707 VI .. KNFALQKIG 717	

FIG. 36
(CONTINUED)

132	LSTQKQYFFQDEWNQVRAMIGNEL . FRHLYTKYLIFQRTSE .. GTLVQFC	178
1	MSRRNQ.....KKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQQI	43
179	GNNVFDHLKVNDKFDDKKQKGGAADMNEPRCCSTCKYNVKNEKDHFLLNNIN	228
44	KEEDLKLLKFKNQDQDGNSGNDDDEE.....NNSNKQQELLRRVN	84
229	VPNWNNMKSRTTRIFYCTHFNRNNQFFKKHEFVSNKNNISAMDRAQTIFTN	278
85QIKQQVQLIKK...VGSKVEKDLNLNEDENKKN	114
279	IIFRFRNIRKKLKDVKIEKIAYMЛЕVKDFNFNYYLTKSCPLPENWRERKQ	328
115	GLSEQQVKEEQLRTITEEQVKYQNLVFNFMDYQLDLNESGGHRRHRRRETDY	164
329	KIENLINKTREEKS SKYYEEELFSYTTDNKCVTQFINE . FFYNILPKDFLTG	377
165	DTEKWFEISHDQK.....NYVSIYANQKTSYCWWLKDYFNK	200
378	RNRKNFQKKVKKYVELNKHELIHKNLLLEKINTREISWMQVETS AKHFY Y	427
201	NNYDHNV SINRLE .. TEAEFYAFDDFSQTIKLTNNSYQTVNID.....	242
428	FDHENIYVLWKLLRWI .. FEDLVVSLIRCFYVTEQQKSYSKTYYRKNI	475
243	VNF DNNLCI ALLRFLLS LERFN ILNIRSSY .. TRNQYNF EKIGELLETI	290
476	WDVIMKMSIADLKETLAEVQEKEVEEWKSLGFAPGKLRLIPKTTFRP	525
291	FAVVFSHR.....HLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQ	330
526	IMTFNKKIVNSDRKTTKLTNTKLLNSHMLMLKTLKNRMFKDPFGFAVFNY	575
331	VYSFSTDLKLV D..TNKVQDYFKFLQEF PRLTHVSQQAIPV SATNAVENL	378

FIG. 37

576	DDVMKKYEEFVCKWKQVGQPKLF.	FATMDIEKCYDS..VNREK	615
379	NVLLKKVKH ANLNLVSIPTQFNDFYFVNQLQHLKLEFGLEPNILTKQK		426
516	LSTFL. KTTKLLSSDFWIMTAQILKRKNNI..VIDSKNFRKEMK		657
427	LENLLLSIKQSKNLKFLRLNFYTYVAQETSRKQILKQATTIKNLKNKNQ		476
558	DYFRQKFQKIALEGQQYPTLFSVLEN..EQNDLNAKKTLIVEAKQRNYFK		705
477	EETPETKDETPSESTSGMKFFDHLSELTELEDFSVN....LQATQEIY..		520
706	KDNLLQPVINICQNYINFNGKFYKQTKGIPQGLCVSSILSSFYyatLEE		755
521	.DSLHKLLIRSTNLKKFKLSSYKYEMEKSMDTFIDLKNI.....YETLNN		564
756	SSLGFLRDESMPNPENPNVNLLMRLTDYLLITTQENNAVLFIEKLINSR		305
565LKRCSVNISNPNGNISYELTN.....KDSTFYKFKLTLNQE		500
806	ENGFKFNMKKLQTSFPLSPSKFAKYGMDSVEEQNIVQDYCDWIGISIDMK		855
601	LQHAKYTFK..QNEFQFNNVKSAKIESSSLESLEDIDS LCKSIASCKNLQ		648
856	TLALMPNINLRIEGILCTLNLMQT..KKASMWLKK..KLKSFLMNNITH		901
649NVNI.....IASLLYPNNIQKNPFPNKPMLLFFKQFEQLKNLENVSINC		691
902	YFRKTI...TTEDFANKTLNKLFISSGGYKYMCAKEYKDHFKKNLAMSSM		948
692	ILDQHILNSISEFLEKNKKIKAFILKRYYLLQYYLDYTKLFKTLQQLPEL		741
949	IDLEVSKIIYSVT.....RAFFKYLVCNIKDT..IFGEEHY		982
742	NQVYINQQLEELTVSEVHKQVWENHKQKAFYEPCLCEFIKESSQTLQLIDF		791
983	PDFFLSTLKHFIIEFSTKLY IFNRVCMILKAKEAKLKS DQCQSLIQ		1028
792	DQNTVSDDS IKKILESISESKYHHYRLRNPSQSSSLIKSENNEIQELLK		840

FIG. 37
(CONTINUED)

4	DIDLDIENLLPNTFNKYSSSCSDKKGCKTLKSGSKSPSLTIPK.....	47
617NVKSAKIESSSLESLEDIDS LCKSIASCKNLQNVNIIASLLYPNNIQKNP	666
48LQKQLEFYFSDANLYNDSFLRKVLVLKSGEQRVE....IETLLM	86
667	FNKPMLLFFKQFEQLKNLENVSINCILDQHILNSISEFLEKNKKIKAFIL	716

FIG. 38

1 MEMDIDLDDIENL....LPNTFMNKYSSSCSDKKGCKTLKSGSKSPS...	42
491 IELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGAKKYGSVRTCLEC	540
43 .LTIPKLQKQ.....LEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLL	85
541 ALVLGLMVKRCEKSSFYIFSSPSSQCNKCYL.EVDLPGDELRPSMQKLL	589

FIG. 39

telomerase p43	LQKOLE FYF SDANLYNDSFLRKLVLKSGEQRVEIETLLM
human La	ICHQUE YYF GD N LP D KFLKEQI.KLDEGWVPLEIMIK
Xenopus LaA	ICEQIE YYF GD N LP D KFLKQQI.LLD D GWVPLETM I K
Drosophila La	ILRO VEYYF GD A NLN R D K FLREQIGKNEDGWVPLSVLVT
S. c. Lhp1p	CLKQ VEFYF SEFNFPYDRFLRTTAEK.NDGWVPISTIAT

FIG. 41

1 aactcatatta attactaatt taatcaacaa gattgataaa aagcagtaaa taaaacccaa
 61 tagatattaat ttagaaagta tcaattgaaa aatggaaatt gaaaacaact aagcacaata
 121 gccaaaaagcc gaaaaattgt ggtggaaact tgaatttagag atgcaagaaa accaaaatga
 181 tatataagtt aggtttaaga ttgacgatcc taagcaatat ctcgtaaacg tcactgcacg
 241 atgtttgtt taggaaggta gttactacta agataaaagat gaaagaagat atatcatcac
 301 taaagcactt cttgaggtgg ctgagtcga tcctgagttc atctgcgtat tggcagtcta
 361 catccgtaat gaactttaca tcagaactac cactaactac attgttagcat ttgtgttgt
 421 ccacaagaat actcaaccat tcatacgaaaa gtacttcaac aaagcagttac ttttgctaa
 481 tgacttactg gaagtctgtg aatttgcata ggttcttat attttgatg caactgaatt
 541 caaaaatttg tatcttgata ggatactttc ataagatatt cgtaaggAAC tcacttccg
 601 taagtgttta caaagatgcg tcagaagcaa gttttctgaa ttcaacaat actaacttg
 661 taagtattgc actgaatcc aacgtaaagaa aacaatgtt cgttactct cagttaccaa
 721 caagtaaaag tgggattaaa ctaagaagaa gagaaaagag aatctctaa ccaaacttta
 781 ggcaataaaag gaatctgaag ataagtccaa gagagaaact ggagacataa tgaacgttga
 841 agatgcaatc aaggctttaa aaccagcagt tatgaagaaa atagccaaga gatagaatgc
 901 catgaagaaa cacatgaagg cacctaaat tcctaactct accttgcata caaagtactt
 961 gaccttcag gatctcatta agttctgcc tatttctgag cttaaagaaa gagttatataa
 1021 gatccttggt aaaaaatacc ctaagaccg agaggaatac aaagcagctt ttggatttc
 1081 tgcattctca cccttcaatc ctgaatttgc tgaaaagcgt atgaagattt aatctctaa
 1141 aacatgggaa aatgaactca gtgc~~aaa~~agg caacactgct gaggttggg ataaatttaat
 1201 ttcaagcaat taactcccatt atatggccat gttacgttaa ttgtcttaca tcttaaaagc
 1261 cggtgttca gatactacac actctattgt gatcaacaag atttgtgagc ccaaggccgt
 1321 tgagaactcc aagatgttcc ctcttcaatt cttagtgcc atgaagactg ttatgaagc
 1381 agttacttaag ggatcaagg ccaagaagag agaaaatatg aatcttaaag gtcaaatcga
 1441 agcgttaaag gaagtgttg aaaaaaccga tgaagagaag aaagatatgg agttggatg
 1501 aaccgaagaa ggagaatttg ttaaagtcaa cgaaggaaat ggcaagaaat acattaactc
 1561 cattgaacctt gcaatcaaga tagcagttaa caagaattta gatgaatca aaggacacac
 1621 tgcaatcttc tctgtatgtt ctggttctat gagtacctca atgtcaggtg gagccaagaa
 1681 gtatggttcc gttcgtactt gtctcgagtg tgcattagtc ctgggttga tggtaaaata
 1741 acgttgtgaa aagtccat tctacatctt cagttcacct agttctcaat gcaataagtgc
 1801 ttacttagaa gttgatctcc ctggagacga actccgtcct tctatgtaaa aacttttgca
 1861 agagaaagga aaacttgggt gtggactga ttccccctat gatgtcattt gatgtggac
 1921 aaaaataaa actcacgttag acaatatcgt tattttgtct gatgtatga ttgcagaagg
 1981 atattcagat atcaatgtta gaggcagttc cattgttaac agcataaaa agtacaaggaa
 2041 tgaagtaat cctaacattt aaatcttgc agttgactta gaaggttacg gaaagtgcct
 2101 taatcttagt gatgagttca atgaaaacaa ctacatcaag atattcggtt tgagcgattc
 2161 aatcttaaag ttcatatccag ccaagcaagg aggagcaat atggtcgaag ttatcaaaaa
 2221 ctttgcctt caaaaatag gacaaaatgt agttcttga gattcttcta taacaaaaat
 2281 ctcacccac tttttgttt tattgcataccattatgaa atttaattt ttatcttattt
 2341 atttaagttt cttacatagt ttatgtatcg cagttatca gcttattcaa atgattctgc
 2401 aaagaacaaa aaagattaaa a

FIG. 42

+

50/103

	Motif A	Motif B	Motif C	Motif D	Motif E
Consensus telomerase p123	h-hDh---h-h	h---+QG---SP	h---YhDDhhh	Gh-h---K	h-hIGh-h
Dong (LINE)	GOPKLFFATMDIEKCYDSVNREKLSTFLKTTKLI-100-KFYKQTKGIPQQGLCVSSIISSSFYYATLEESSLGFLL	KNRNLHCTYDDYKKAFDSIPHSMWLIQVLEIYKIN-	-14-LMRLTDDYLLITTCQENN-0-AVLFIEKLINVSRENGFKENMKKLQT-23-QDYCDWIGISI	-16-HIYMDIDKLYAKNDKE-0-MKKLIDDTTIFNSNDISMQFGIDKCKT-25-KCLYKYLGFQQ	
a1 S.c. (groupII)	FGGSNWFERVDLKKCFDTISHDLIKELKRYISD-	LKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTTIP-	-55-YVRYADDILIGVUGSKN-2-KIKRDLNFLNS.LGLTINEBEKTLLI-4-ETPARFLGYNI	-4-IYQYMDLIVGSHLEIG-1-HRTKIEELROHLLRWGLTPDKXKHOK-0-EPPFLWMGYEL	
HIV-RT	VLPELYFMKFDVKSCYDSIPRMECMRILKDALKN-	68-KCYTREDGGLFQGSSLSSAPIVDLVYDDLLFEYSEFK	-8-ILKLAADDFLIISTDQQQ.....VINIKKLAMGGFQKYNAKANR-41-IRSKSSKGIFR		
L8543 .12					

FIG. 40

+

MEIENNQAQQPKAEKLWWLELEMQENQNDIQVRVKIDDPKQYL
VNVTAAACLQLQEGSYYQDKDERRYIITKALLEVAESDPEFICQLA
VYIRNELYIRTNTYIVAFCVHKNTQPFIIEKYFNKAVALLPNDL
LEVCEFAQVLYIFDATEFKNLYLDRILSQDIRKELTFRKCLQR
VRSKFSEFFNEYQLGKYCTESQRKKTMFRLSVTNKQWDQTKK
RKENLLTKLQAIKESEDKSKEKGDIIMNVEDAIKALPAVMKK
AKRQNAMKKHMKAQKIPNSTLESKYLTFKDLIKFCHISEP
YKLIGKKYPKTEEEYKAAGDSASAPFNPELAGKRMKIEISKT
ENELSAKGNTAEVWDNLISNQLPYMAMLRNLSNIKAGVSDT
HSIVINKICEPKAVENSKMFPQLQFFSAIEAVNEAVTKGFKA
ENMNLLKGQIEAVKEVVEKTDEEKKDMELEQTEEGEFVKVNEG
GIGQYINSIELAIKIAVNKNLDEIKGHTAIFSDVSGSMS
MSGGA
KKYGSVRTCLECALVLGLMVQRC
EVDSLPGDELPSMQKLLQEKGKLGGGDFPYECIDE
EVNIVILSDMMIAEGYSDINVRGSSIVNSIKKYK
DEVNPNIKIF
VDLEGYGCNLNGDEFNENNYIKIFGMSDSILKF
FISAKQGGANM
VEVIKNFALQKIGQK

FIG. 43

MSRRNQKKPQAPIGNETNLDLDFVLQNLEVYKSQIEHYKTQQQQIK
EEDLKLLKFKNQDQDGNSGNDDEENNSNKQQELLRRVNQIKQ
QVQLIKKVGSKVEKDLNLNEDENKKNGLSEQQVKEEQLRTITEE
QVKYQNLVFNMDDYQQLDNLNESGGHRRRERTDYDTEKFWEISHDQ
KNYVSUYANQKTSYCWWLKDYNKNNYDHLDLNVSINRLETAEYF
AFDDFSQTIKLNTNSYQTVNIDVNFDNNCLILALLRFLLSLERF
NILNRSSYTRNQYNFKEKIGELLETIFAVVFSHRHLQGIHLQVP
CEAFQYLVNSSSQISVKDSQLQVYSFSTDLKLVDTNKVQDYFKF
LQEFPRLTHVSQQAIPVSATNAVENLNVLKKVKHANLNLVSIPI
TQFNFDYFVNLQHLLKEFGLEPNILTQKQLENLLS1KQSKNL
KFLRLNFTYVAQETSRKQ1LKQATTI1CNLKNNKQNEETPETKD
ETPSESTSGMKFFDHLSELTELEDFSVNLQATQEIYDLSHLKL1
RSTNLKKFKLSQLSYKEMEKSMDTFIDLKN1YETLNNLKRC3VNI
SNPHGNISYELTNKDSTFYKFKLTLNQELQHAKYTFKQNEFQFN
NVSAKIESSSLESLEDIDS1LCKSIASCKNLQNVNIIASLLYPN
NIQKNPFNPKPNLLFFKQFEQLKNLENVSINCILDQHILNS1SEF
LEKNNKK1KAFLKRYYLLQYYLDYTKLFKTLQQLPELNQVYINQ
QLEELTVEVHKQWENHDKQAFYEP1LCEFIKESSQTLQLIDFD
QNTVSDDS1KK1LIES1SESKYH1LRLNPSPQS11KSENEEITQ
ELLKACDEKGVLVKAYYKFPLCLPTGTYYDYNNSDRW

FIG. 45

MKILFEFIQDKLDidLQTNSTYKENLKCGHFnGLDEI LTTCFAL
PNSRKIALPCPGDLSHKAVIDHCIIYLLTGELYNNVLTFGYKI
ARNEDVNNSLFCHSANVNVTLLKGAAWKMfhSLVGTyAFVDLLI
NYTVIQFNGQFFTQIVGNRCNEPHLPPKwVQRSSSSATAAQIK
QLTEPVTKQFLHKLNINSSSSFPYSKILPSSSSIKLTDLRLA
I FPTNLVUKI PQRKVKRINLTLQKLLKRHKRLNIVSILNSICPLP
EGTVLDLShLRSQPSPKERVLFKIIIVLQKLLPQEMFGSKKNKGK
I IKNLNLLSLPLNGYLPFDSSLKKLRKDFRWLFTISDIWFTHH
NFENLNQLAICFISWLFRQLIPKII QTFFYCTEISSTVTIVYFR
HDTWNKLITPFIVEYFKTYLVEENNVCRNHNHSYTLSNFNHSKMR
I IPKKSNEFRIIAI PCRGADEEEFITYKENHKNAI QPTQKILEY
LRNKRPTSTFKIYSPQTQIADRIKEFKQRLKKFNNVLPELYFMK
FDVKSCYDSIPRMCEMCRILKDALKNENGFFVRSQYFFNTNTGVL
KLFNVNVNASRVPKPVEYLIDNVRTVHLSNQDVINVVEMEIFKTA
LWVEDKCYIREDGLFOGSSLSAPIVDLVYDDLLEFYSEFKASPS
QDTLILKLADDFLIISTDQQQVINNIKKLAMGGFQKYNAKANRDK
ILAVSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMNFHIRSKS
SKGIFRSLIALFNTRISYKTIDTNLNSTNTVLMQIDHVVKNISE
CYKSAFKDLSINVTONMQFHSFLQRRIIEMTVSGCPITKCDPLIE
YEVRFTILINGFLESLSNTSKFDNIIILLRKEIQHLQAYIYIYI
HIVN

FIG. 46

FIG. 44

+

53/103

Oxytricha
Euplotes

LCVSYILSSFYANLEENALQFLRKEAMDPEKPETNLLMRLT
LCVSSILSSFYATLEESSLGFLRDESMNPENPNVNLLMRLT

FIG. 47

ATTTATACTCATGAAAATCTTATTGAGTTCAAGACAAGCTGACATTGATCTACA
GACCAACAGTACTTACAAAGAAAATTAAAATGTGGTCACCTCAATGGCCTCGATGAAAT
TCTAACTACGTGTTCGCACTACCAAATTCAAGAAAATAGCATTACCATGCCTTCCTGG
TGACTTAAGCCACAAAGCAGTCATTGATCACTGCATCATTACCTGTTGACGGCGAATT
ATACAACAAACGTACTAACATTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAATAG
TCTTTTTGCCATTCTGAAATGTTAACGTTACGTTACTGAAAGGCCGTGCTGGAAAAT
GTTCCACAGTTGGTCGGTACATACGCATTGTTGATTATTGATCAATTATACAGTAAT
TCAATTAAATGGGCAGTTTCACTCAAATCGGGTAACAGATGTAACGAACCTCATCT
GCCGCCAAATGGGCTAACGATCATCCTCATCATCCGCAACTGCTGCGCAAATCAAACAA
ACTTACAGAACCGAGTACAATAAACATTCTACACAAGCTCAATATAAATTCCCTTTC
TTTTTTCTTATAGCAAGATCCTCCTCATCATCATCTATCAAAAAGCTAACTGACTT
GAGAGAAGCTATTTTCCCACAAATTGGTAAAATTCCCTAGAGACTAAAGGTACGAAT
TAATTGACGCTGCAAAGCTATTAAAGAGACATAAGCGTTGAATTACGTTCTATTT
GAATAGTATTGCCACCATTGAAAGGGACCGTATTGGACTTGTGCGATTGAGTAGGCA
ATCACCAAAGGAACGAGTCTGAAATTATCATTGTTATTACAGAAGTTATTACCCCA
AGAAATGTTGGCTCAAAGAAAATAAGAAAATTATCAGAATCTAAATCTTTATT
AAGTTTACCTTAAAGGCTATTACCATTTGATAGTTGTTGAAAAGTTAAAGATTAA
GGATTTCGGTGGTTGTTCATTTCTGATATTGGTTCACCAAGCACAATTGGAAAATT
GAATCAATTGGCATTGTTCATTTCTGGTATTAGACAACATAATTCCCAAATT
ACAGACTTTTTTACTGCACCGAAATATCTTCTACAGTGACAATTGTTACTTTAGACA
TGATACTTGGAAATAAACTTATCACCCCTTTATCGTAGAATATTAAAGACGTACTTAGT
CGAAAACAACGTATGAGAACATAATAGTTACACGTTGCAATTCAATCATAGCAA
AATGAGGATTATACAAAAAAAGTAATAATGAGTTCAAGGATTATTGCCATCCATGCAG
AGGGGCAGACGAAGAAGAATTCAAAATTATAAGGAGAACACAAAAATGCTATCCAGCC
CACTAAAAAAATTAGAATACCTAAGAAACAAAGGCCACTAGTTTACTAAAATATA
TTCTCCAACGAAATAGCTGACCGTATCAAAGAATTAAAGCAGAGACTTTAAAGAAATT
TAATAATGTCTTACAGAGCTTATTTCATGAAATTGATGTCAAATCTGCTATGATT
CATACCAAGGATGGAATGTATGAGGAACTCAAGGATGCGCTAAAAAATGAAAATGGTT
TTTCGTTAGATCTCAATATTCTCAATACCAATACAGGTGATTGAAGTTATTAAATGT
TGTAAACGCTAGCAGAGTACCAAAACCTTATGAGCTATACATAGATAATGTGAGGACGGT
TCATTTATCAAATCAGGATGTTATAACGTTAGAGATGAAATATTAAAACAGCTT
GTGGGTTGAAGATAAGTGTACATTAGAGAAGATGGCTTTTCAGGGCTCTAGTTTATC
TGCTCCGATCGTGTATTGGTGTATGACGATCTCTGGAGTTTATAGCGAGTTAAAGC
CAGTCTCTAGCCAGGACACATTAATTAAAAGCTGGCTGACGATTCTTCTATAATATCAAC
AGACCAACAGCAAGTGTCAATATCAAAGCTGGCATGGGGATTTCAAAATATAA
TGCAGAAAGCCAATAGAGACAAAATTAGCCGTAAGCTCCAAATCAGATGATGATACGGT
TATTCAATTGTGCAATGCACATATTGTTAAAGAATTGGAAGTTGGAAACATTCAAG
ACAATGAATAATTCCATATCGTTGAAATCTAGTAAAGGGATATTGCAAGTTAAAT
AGCGCTGTTAACACTAGAATCTCTTATAAAACAATTGACACAAATTAAATTCAACAAA
CACCGTTCTCATGCAAATTGATCATGTTGAAAGAACATTGCGAATTGCAATT
TTTTAAGGATCTATCAATTAAATGTTACGAAAATATGCAATTGCAATTGTTACACG
CATCATTGAAATGACAGTCAGCGGTTGCAATTACGAAATGTGATCCTTAAATCGAGTA
TGAGGTACGATTACCCATATTGAATGGATTGGAAAGCCTATCTCAAACACATCAA
ATTTAAAGATAATATCATTCTTGAGAAAGGAAATTCAACACTTGCAAGC

FIG. 48

+

+

54/103

AKFLHWLMSVYVVELRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKR
VQLRDVSEAEVRQHREARPALLSRLRFIPKPDGLRPIVNMDYVGARTFRREK
RAERLTSRVKALFSVLNYERA

FIG. 49

GCCAAGTTCCCTGCACTGGCTGATGAGTGTACGTCGAGCTGCTCAGGTC
TTCTTTTATGTCACGGAGACCACGTTCAAAAGAACAGGCTCTTTCTACC
GGAAGAGTGTCTGGAGCAAGTTCAAAGCATTGGAATCAGACAGCACTTGAG
AGGGTGCAAGCTGCAGCTCGGAAGCAGAGGTCAAGGCAGCATCGGGAAAGC
CAGGCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACGGGC
TGCAGGCCGATTGTGAAACATGGACTACGTGTTGGAGCCAGAACGTTCCGAGA
GAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTAGCGTGCT
CAAATACGAGCGGGCGCG

FIG. 50

MTEHHTPKSRILRFLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLSDVQTSFSIFLHSTVVG
DSKPDEGVQFSSPKCSQSELIANVVQMFDESFRERRNLLMKGFSMNEDFRAMHNGVQNDLVSTF
PNYLISILESKNWQLLIEIIGSDAMHYLLSKGSIFEALPNDNYLQISIGIPLFKNNVFEETVSKKRKR
TIETSITQNKSRKEVSWNSISISRFSIFYRSSYKKFKQDLYFNLHSICDRNTVHMWLQWIFPRQFG
LINAQVKQLHKVIPLVSQSTVVPKRLVKVYPLIEQTAKRLHRISLSKVYNHYCPYIDTHDDEKILS
YSLKPNQVFAFRLSILVRVFPKLIWGNQRIFEIILKDLETFLKLSRYESFSLHYLMSNIKISEIEWL
VLGKRSNAKMCLSDFEKRQIIFAIFIYWLNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCR
PFITSMKMEAFEKINENNVRMDTQKTLPPAVIRLLPKKNTFRLITNLRKRFLIKMGSNKKMLVSTN
QTLRPVASILKHLINEESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDI
KSCYDRIKQDLM
FRIVKKKLDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPEKVVQLLSMKTSDTLFVDFV
WTKSSEIFKMLKEHLSGHIVKIGNSQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL
LRRVDDFLFITVNKKDAKKFLNLSLRGFKEHNFSTSLEKTVINFENSNGIINNTFFNESKKRMPFFG
FSVNMRSLDTLLACPKIDEALFNSTSVELTKHMGKSFYKILRSSLASFAQFIDITHNSKFN
IYRLGYSMCMRAQAYLKRMDIFIPQRMFITDLLNVIGRKIWKKLAEILGYTSRRFLSSAEVKWLFC
LGMRDGLKPSFKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVLFLHRR
IAD

FIG. 51

+

FIG. 52

FIG. 52
(CONTINUED)

57/103

FIG. 53

+

58/103

S-1: FFY VTE TTF QKN RLF FYR KSV WSK
S-2: RQH LKR VQL RDV SEA EVR QHR EA
S-3: ART FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q
A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS
A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

FIG. 54

Poly 4

t t c
t a a g c c t c g
5' - c a g a c c a a a g g a a t t c c a t a a g g - 3'
Q T K G I P Q G

4 (B')

5 (c')

D D Y L L I T
3' - c t g c t g a t g g a g g a g a t a g t g g - 5'
a a a a a a a a
t t t t

Poly 1

FIG. 56

+

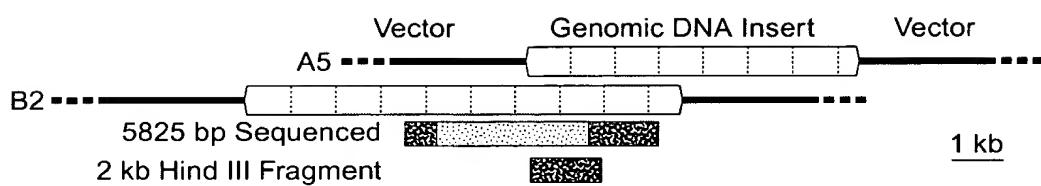


FIG. 55A

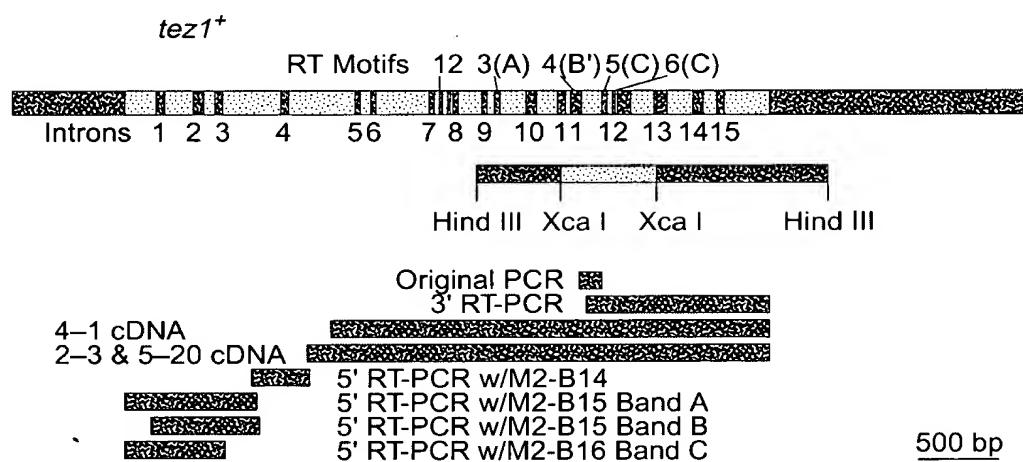


FIG. 55B

+

60/103

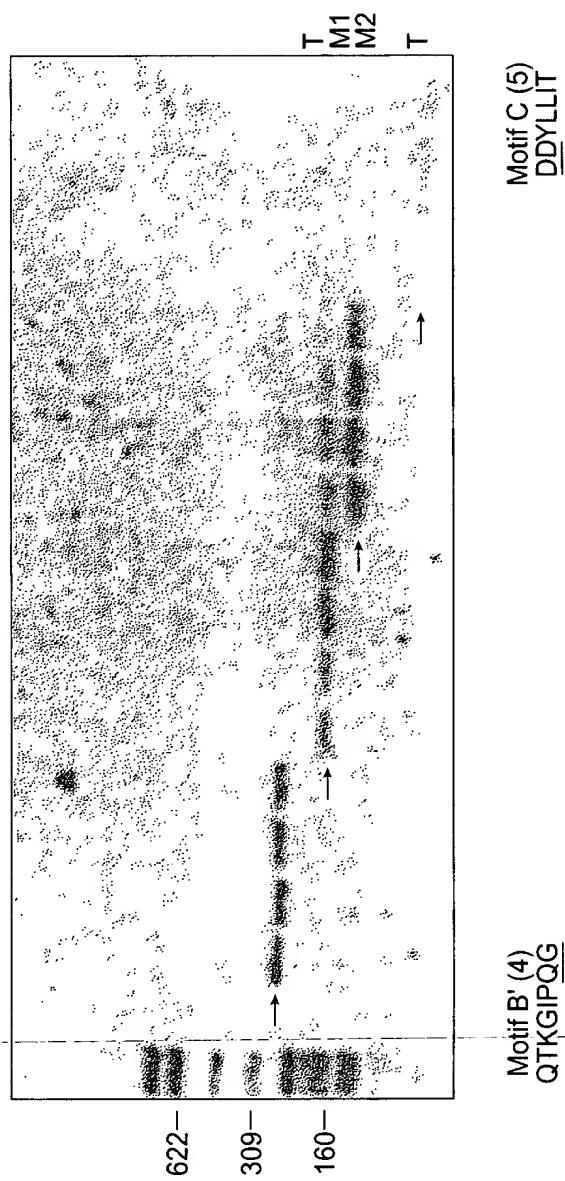


FIG. 57

+

+

61/103

Ot	LCVSYILSSFYANLEENALQFLRKESMDPEKPKPETNLLMRLT
Ea_p123	KGIPQGLCVSSILSSFYATLEESSLGFLRDESMNPNPVNVLMLRLTDDYLLIT
Sp_M2	SILSSFLCHFYMEDLIDEYLTSFTKK-----GSVLLRVV
Sc_p103	DGLFQGSSLSAPIVDLVYDDLEFYSEFKASPS-----QDTLILKIAADDFLIIS
	* . * . *
Q K V G I P Q G	<----Actual Genomic Sequence.
caa aaa gtt ggt atc cct cag gag.....	
<u>Poly 4</u>	
t a a g	t c t c g
cag acc aaa gga	att cca taa gg ----->
ag acc aaa gga	att cca tca ggc tca ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG
tc tgg ttt cct	cct taa ggt agt ccg AGT TAA GAC AGT AGA AAA AAC ACA GTA AAG ATA TAC
K G I P S G S I L S S F L C H F Y M	

FIG. 58

+

+

62/103

GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA TCA GTG TTG TTA CGA
CTT CTA AAC CTA CTT ATG GAT AGC AAA TGC TTT TTC CCT AGT CAC AAC AAT GCT
E D L I D E Y L S F T K K G S V L L R

GTA GTC gac tac ctc ctc atc acc
CAT CAG ctg ctg atg gag gag tag tgg

V V D D Y L L I T

<---- ctg ctg atg gag gag tag tgg
a a a a a a a a

t t t t t t

c c

Poly 1

.....gac gat ttc ctc ttt ata aca.....
D D F L F I T <---Actual Genomic Sequence

FIG. 58
(CONTINUED)

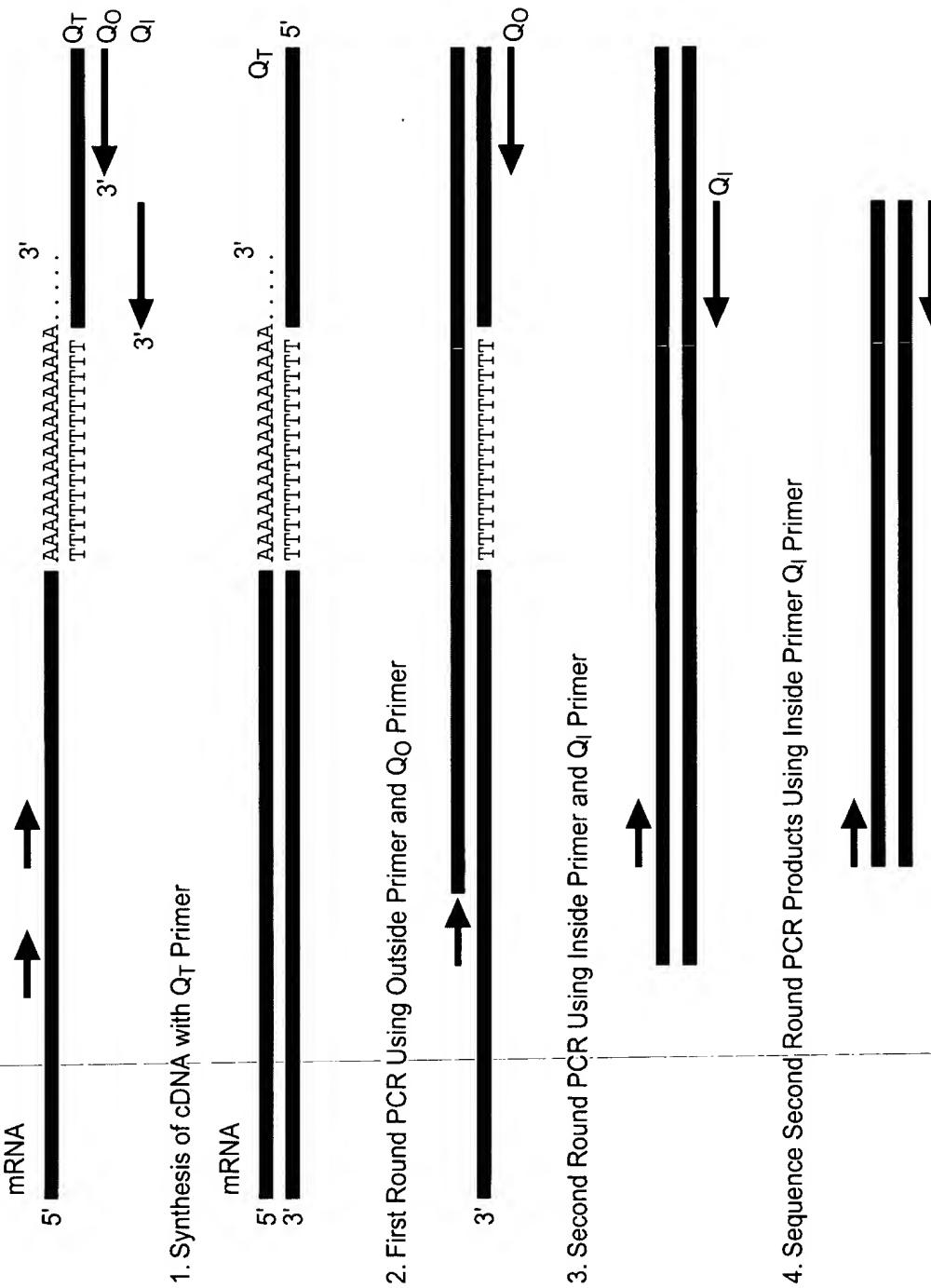


FIG. 59

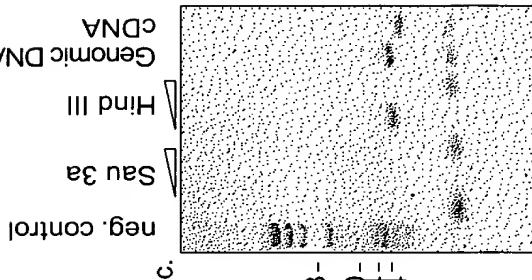
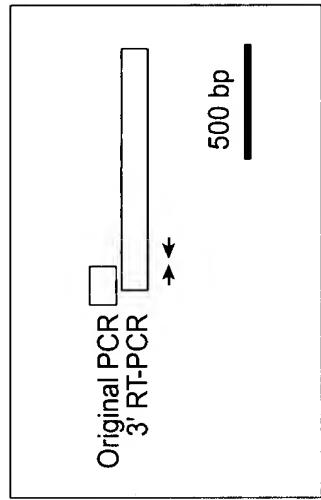
A. Genomic Libraries
Size Selected Libraries from *P. Nurse*

3~4 kb
5~7 kb
7~8 kb
11~12 kb

Libraries from J.A. Wise
Sau 3a Partial Digest
Hind III Partial Digest

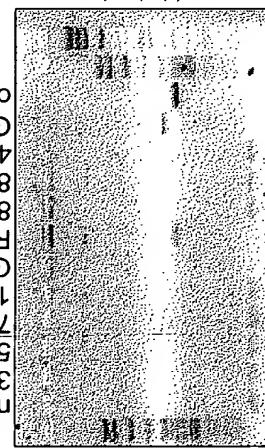
cDNA Libraries
GAD (Gal Activation Domain) Library
REP Library from R. Allshire
REP81ES Library (old)
REP81ES Library (new)
REP41ES Library

B.



D. Mg. Conc.

P. Nurse cDNA Lib
GAD cDNA Lib
REP Lib from Robin
81ES Lib Old
81ES Lib New
41ES Lib New
Genomic DNA



C.

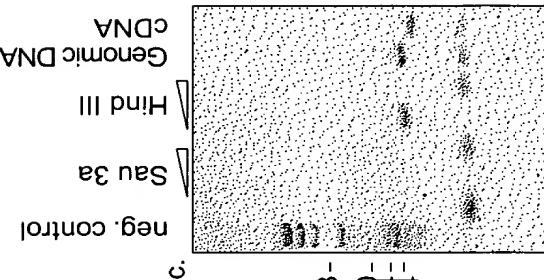


FIG. 60

+

65/103

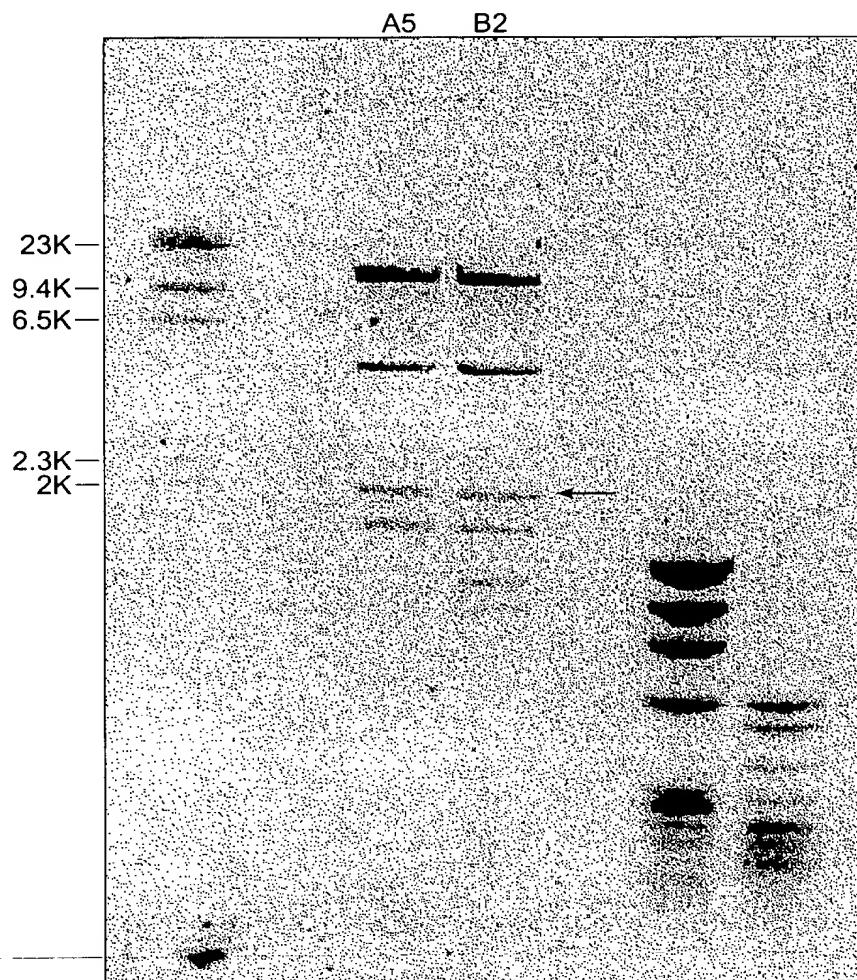
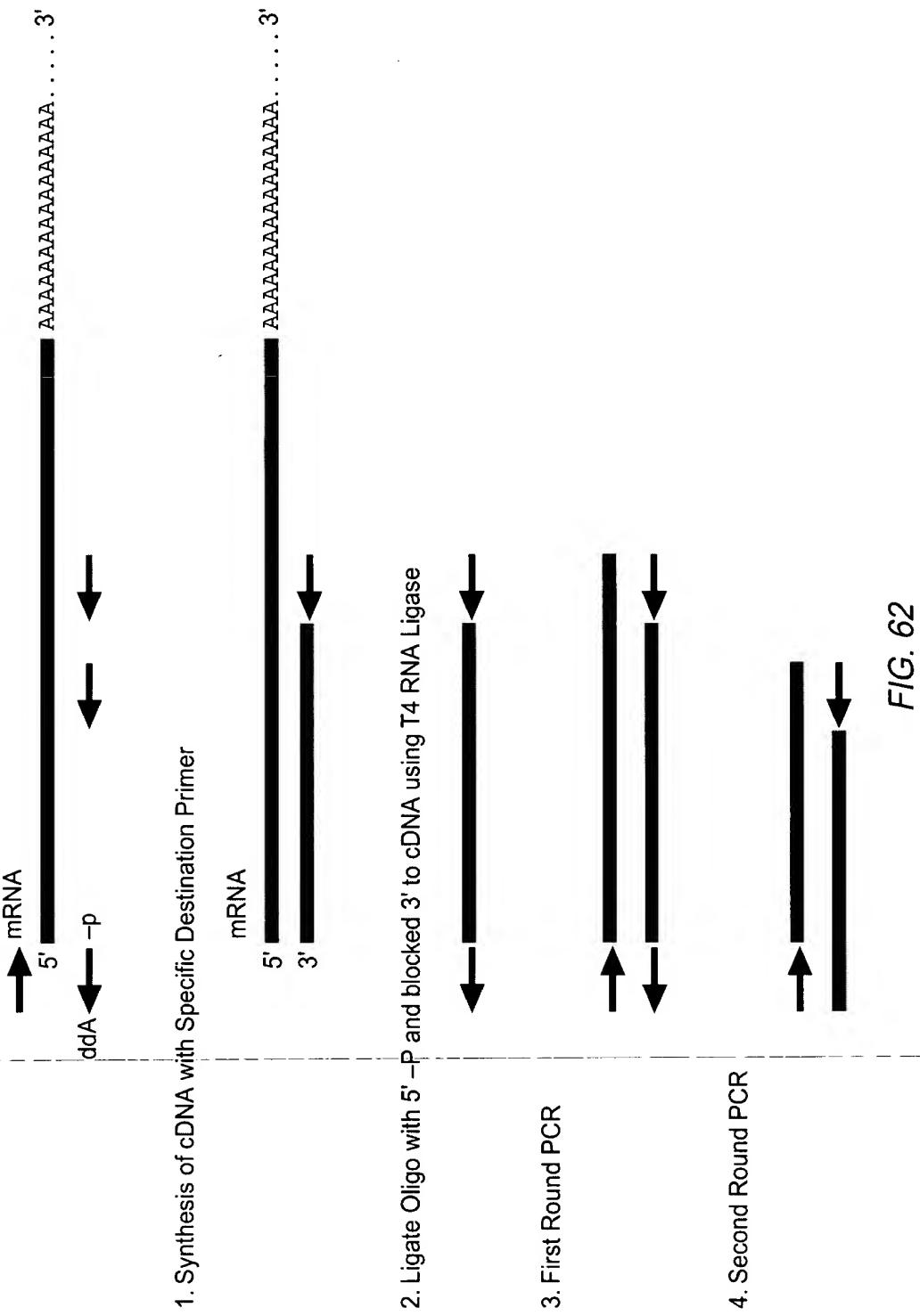


FIG. 61

+



				Motif 0
S.p.	Tez1p	(429)	. WLYNSFIIPILQSFYITTESSDLRNRTVYFRKDIW	... (35) ...
S.c.	Est2p	(366)	. WLFRQLIPKIIQTFYYCITEISSTVT-IVYFRHDTW	... (35) ...
E.a.	p123	(441)	. WIFEDLVVSLIRCFYYVTEQQSYSKTYYRKNIW	... (35) ...
		*	*** * *	*
		Motif 1	Motif 2	K
		p hh h	hR h	R
S.p.	Tez1p	AVIRLLPKK-	-NTFLRLITN-LRKRF	... (61) ...
S.c.	Est2p	SKMRIIPKKSNNEFRIIAIPCRGAD	... (62) ...	
E.a.	p123	GKRLIPKK-	-TTERPIMTFNKIV	... (61) ...
		*	***	*
		Motif 3 (A) AF		
		h hDh	GY	h
S.p.	Tez1p	KKYFVRIDIKSCYDRIKODLMERIVK	... (89) ...	
S.c.	Est2p	ELYFMKFDVKSCYDSIIPRMECMRLK	... (75) ...	
E.a.	p123	KLFFATMDIEKCYDSVNREKLSTFLK	... (107) ...	
		*	***	*
		Motif 4 (B')		
		hPQG	PP hh	h
S.p.	Tez1p	YLQKVGTIPQGSILSSFLCHFYMEDLIDEYLSF	... (6) ...	
S.c.	Est2p	YIREDGFLFGQSSLSSAPIVDLVYDDLFYSEF	... (8) ...	
E.a.	p123	YKQTKGIPQGLCVSSILSSFFYYATLEESSLGF	... (14) ...	
		*	**	*
		Y Motif 5 (C)		
		h F DDhh	Gh CK	h
S.p.	Tez1p	VLLRVVDDFLFITVNKKDAKKFLNLSLRGEKHNFSTSLEKTIVINFENS	... (205)	
S.c.	Est2p	LILKLAADDFLIIISTDQQVINIKKLAMGGFQYNAKANRDKILAVSSQS	... (173)	
E.a.	p123	LLMRLTDDYLLITQTENNNAVLFIEKLINVSRENGFKPNMKLQLTSFPLS	... (209)	
		***	***	*

FIG. 63

A.

Sp_Tip1p	1	- - - - -	M T E H H T P K S R I L R F L E N Q Y V Y L C T	24
Sc_Est2p	1	- - - - -	M K I L F E F	7
Ea_p123	1	M E V D V D N Q A D N H G I H S A L K T C E E I K E A K T L Y S W	33	
Sp_Tip1p	25	L N D Y V Q L V [LRG SPA S S Y S N I C E R L R S D V Q T S F S	57	
Sc_Est2p	8	I Q D K L D I D L Q T N - - S T Y K - - E N L K C G H F N G L D	35	
Ea_p123	34	I Q K V I R C R N Q S Q - - S H Y K - - D L E D I K I F A Q T N	61	
Sp_Tip1p	58	I F L H S T V V G F D S K P D E G V Q F S S P K C S Q S E L I A N	90	
Sc_Est2p	36	E I L T T C F A L P N S R - K I A L P C L P G D L S H K A V I D H	67	
Ea_p123	62	I V A T P R D Y N E E D F K V I A R K E V F S T G L M I E L I D K	94	
Sp_Tip1p	91	V V K Q M F D E S F E R R R - N L L M K G F S M N H E D F R A M H	122	
Sc_Est2p	68	C I Y L L T G E L Y N - - N V L T F G Y K I A R N E D - - -	93	
Ea_p123	95	C L V E L L S S S D V S D R Q K L Q C F G F Q L K G N Q - - -	122	
Sp_Tip1p	123	V N G V Q N D L V S T F P N Y L I S I L E S K N W Q L L E I G	155	
Sc_Est2p	94	- - V N N S L F C H S A N V N V T L L K G A A W K M F H S L V G	123	
Ea_p123	123	- - L A K T H L L T A L S T Q K Q Y F F Q D E W N Q V R A M I G	152	
Sp_Tip1p	156	S D A I M H Y L L S K G S I F E A L P N D N Y L Q I S G I P L F K N	188	
Sc_Est2p	124	T Y A F V D L L I N Y T V I Q F N - G Q F F T Q I V G N R C N E P	155	
Ea_p123	153	N E L F R H L Y T K Y L I F Q R T S E G T L V Q F C G N N V F D H	185	
Sp_Tip1p	189	N V F E E T V S K K R K R T I E T S I T Q N - - - K S A R K E V S	218	
Sc_Est2p	156	H L P P K W V Q - R S S S S A T A A Q I - - - K Q L T E P V T	183	
Ea_p123	186	L K V N D K F D K - K Q K G G A A D M N E P R C C S T C K Y N V K	217	

68/103

FIG. 64

+

+

69/103

A.

Sp_Tip1p	219	WN S I S I S R F S I F Y R S S Y K K F K Q D L Y F N L H S I C D	251
Sc_Est2p	184	N - - - - -	200
Ea_p123	218	N E K - - - - -	248
Sp_Tip1p	252	R N T V H M W L Q W I F P R Q F G L I N A F Q V K Q L H K V I P L	284
Sc_Est2p	201	R - - - - -	223
Ea_p123	249	R - - - - -	275
Sp_Tip1p	285	V S - - - Q S T Y V P K R L L K V Y P L I E Q T A K R L H R I S	313
Sc_Est2p	224	T N - - - L V K I P Q R L K V R I N L T L Q K L L K R H K R L N	252
Ea_p123	276	F T N I F R F N R I R K K L K D K V I E K I A Y M L E K V K D F N	308
Sp_Tip1p	314	L S K V Y N H Y C P Y I D T H D D E K I L S Y S L K P N Q - -	342
Sc_Est2p	253	Y V S I L N S I C P P L E G T V D L S H L S R Q S P K E R - -	282
Ea_p123	309	F N Y Y L T K S C P L P E N W R E K Q K I E N L I N K T R E E K	341
Sp_Tip1p	343	- - - - - - - - - - -	359
Sc_Est2p	283	- - - - - - - - - - -	299
Ea_p123	342	S K Y Y E E L F S Y T T D N K C V T Q F I N E F F Y N I L P K D F	374
Sp_Tip1p	360	W G N Q R I F E I I L K D L E T F L K L S R Y E S F S L H Y L M S	392
Sc_Est2p	300	F G S K K N K G K I I K N L N L L S L P L N G Y L P F D S L L K	332
Ea_p123	375	L T G - R N R K N F Q K K V K K Y V E L N K H E L I H K N L L E	406
Sp_Tip1p	393	N I K I S E I E W L V L G K R S N A K M C L S D F E K R K Q I F A	425
Sc_Est2p	333	K L R I K D F R W L F I S - - D I W F T K H N F E N L N Q L A I	362
Ea_p123	407	K I N T R E I S W M Q V E T S - A K H F Y Y F D H E N - I Y V L W	437

FIG. 64
(CONTINUED)

+

+

70/103

A.

Sp_Tip1p	426	E F I Y W L Y N S F I I P I L Q S F F Y I T E S S D L R N R T V Y	458
Sc_Est2p	363	C F I S W L F R Q L I P K I I Q T F F Y C T E I S S T V T - I V Y	394
Ea_p123	438	K L L R W I F E D L V V S L I R C F F Y V T E Q Q K S Y S K T Y Y	470
Sp_Tip1p	459	F R K D I W K L L C R P F I T S M K M E A F E K I N E N N V R M D	491
Sc_Est2p	395	F R H D T W N K L I T P F I V E Y F K T Y L V E N N V C R N H N S	427
Ea_p123	471	Y R K N I W D V I M K M S I A D L K K E T L A E V Q E K E V E E W	503
Sp_Tip1p	492	T Q K T T L P P A V I R L L P K K - N T F R L I T N L R K R F L	522
Sc_Est2p	428	Y T L S N F N H S K M R I I P K K S N N E F R I I A I P C R G A D	460
Ea_p123	504	K K S L G F A P G K L R L I P K K - T T F R P I M T F N K K I V	534
Sp_Tip1p	523	I K M G S N K K M L V S T N Q T L R P V A S I L K H L I N E - -	552
Sc_Est2p	461	E E E - - F T I Y K E N H K N A I Q P T Q K I L E Y L R N K R P T	491
Ea_p123	535	N S D - - R K T T K L T T N T K L L N S H L M L K T L K N R - M F	564
Sp_Tip1p	553	E S S G I P F N L E V Y M K L L T F K K D L L K H R M F G R - K K	584
Sc_Est2p	492	S F T K I Y S P T Q I A D R I K E F K Q R L L K F N N V L P E L	524
Ea_p123	565	K D P F G F A V F N Y D D V M K K Y E E F V C K W K Q V G Q P K L	597
Sp_Tip1p	585	Y F V R I D I K S C Y D R I K Q D L M F R I V K K K L K D P E - F	616
Sc_Est2p	525	Y F M K F D V K S C Y D S I P R M E C M R I L K D A L K N E N G F	557
Ea_p123	598	F F A T M D I E K C Y D S V N R E K L S T F L K T T K L L S S D F	630
Sp_Tip1p	617	V I R K Y A T I H A T S D R A T K N - - - - -	634
Sc_Est2p	558	F V R S Q Y F F N T N T G - - - - -	570
Ea_p123	631	W I M T A Q I L K R K N N I V I D S K N F R K K E M K D Y F R Q K	663

FIG. 64
(CONTINUED)

+

A.	Sp_Tip1p	635	F V S E A F S Y F D M V P F E K V V Q L L S - - M K T S D T L F V	665
	Sc_Est2p	571	- - - - - V L K I F N V V N A S R - - V P K P Y E L Y -	591
	Ea_p123	664	F Q K I A L E G G Q Y P T L F S V L E N E Q N D L N A K K T L I V	696
Sp_Tip1p	666	D F V D Y W T K S S S E I F K M L K E H L S G H I V K I G N S Q Y	698	
Sc_Est2p	592	D N V R T V H L S N Q D V I N V V E M E I F K T A L W V E D K C Y	624	
Ea_p123	697	E A K Q R N Y F K K D N L L Q P V I N I C Q Y N Y I N F N G K F Y	729	
Sp_Tip1p	699	L Q K V G I P Q G S I L S S F L C H F Y M E D L I D E Y L S F T K	731	
Sc_Est2p	625	I R E D G L F Q G S S L S A P I V D L V Y D D L L E F Y S E F K A	657	
Ea_p123	730	K Q T K G I P Q Q G L C V S S I L S S F Y Y A T L E E S S L G F L R	762	
Sp_Tip1p	732	K K G - - - - S V L L R V V D D F L F I T V N K K D A K K	756	
Sc_Est2p	658	S P S Q D - - - - T L I J K L A D D F L I I S T D Q Q Q V I N	684	
Ea_p123	763	D E S M N P E N P N V N I L M R L T D D Y L L I T T Q E N N A V L	795	
Sp_Tip1p	757	F L N L S L R G F E K H N F S T S L E K T V I N F E N S N G - -	786	
Sc_Est2p	685	- K K L A M G G F Q K Y N A K A N R D K I L A V S S Q S D - -	713	
Ea_p123	796	F I E K L I N V S R E N G F K F N M K K L Q T S F P L S P S K F A	828	
Sp_Tip1p	787	- - I I N N T F F N E S K K R M P F F G F S V N M R S L D T L L	816	
Sc_Est2p	714	- D D T V I Q F C A - M H I F V K E L E V W K H S S T M	739	
Ea_p123	829	K Y G M D S V E E Q N I V Q D Y C D W I G I S I D M K T L A L M P	861	
Sp_Tip1p	817	A C P K I D E A L F N S T S V E L T K H M G K S F F Y K I L R S S	849	
Sc_Est2p	740	N N F H I R S K S S K G I F R S L I A L F N T R I S Y K T I D T N	772	
Ea_p123	862	N I N L R I E G I L C T L N L N M Q T K K A S M W L K K K L K S F	894	

FIG. 64
(CONTINUED)

A.								
Sp_Tip1p	850	L A S F A Q V F I D I T H N S K F N S C C N I Y R L G Y S M C M R	882					
Sc_Est2p	773	L N S T N T V L M Q I D H V V K N I S E C - - - - -	793					
Ea_p123	895	L M N N I T H Y F R K T I T T E D F A N K T L N K L F I S G G Y K	927					
Sp_Tip1p	883	A Q A Y L K R M K D I F I P Q R M F I T D L L N V I G R K I W K K	915					
Sc_Est2p	794	- - - Y K S A F K D L S I N - V T Q N M Q F H S F L Q R I E M	821					
Ea_p123	928	Y M Q C A K E Y K D H F K K N L A M S S M I D L E V S K I I Y S V	960					
Sp_Tip1p	916	L A E I I L G Y T S R R F L S S A E V K W L F C L G M R D G L K P S	948					
Sc_Est2p	822	T V S G C P I T K C D P L I E Y E V R F T I L N G F L E S L S S N	854					
Ea_p123	961	T R A F F K Y L V C N I K D T I F G E E H Y P D F F L S T L K H F	993					
Sp_Tip1p	949	F K Y H P C F E Q L I Y Q F Q S L T D I I K P L R P V L R Q V L F	981					
Sc_Est2p	855	T S - - - - - K F K D N I I I L R K E I Q H L Q A Y I Y	877					
Ea_p123	994	I E I F S - - - T K K Y I E N R V C M I I L K A K E A K L K S D Q C	1023					
Sp_Tip1p	982	L H R R I A D -	988					
Sc_Est2p	878	I Y I H I V N -	884					
Ea_p123	1024	Q S L I Q Y D A	1031					

FIG. 64
(CONTINUED)

73/103

B.	Sp_Tip1p	1	- - - - -	M T E H H T P K S R I L R F L E N Q Y V Y L C T	²⁴
	Sc_Est2p	1	- - - - -	- M K I L F E F	7
	EA_p123	1	ME V D V D N Q A D N H G I H S A L K T C E E I K E A K T L Y S W	33	
Sp_Tip1p	25	L N D Y V Q L V L R G S P A [S S Y S N I C E R I L R S D V Q T S F S	57		
Sc_Est2p	8	I Q D K L D I D L Q T N - - S T Y K - - EN L K C G H F N G L D	35		
EA_p123	34	I Q K V I R C R N Q S Q - - S H Y K - - D L E D I K I F A Q T N	61		
Sp_Tip1p	58	I F L H S T V V G F D S K P D E G V Q F S S P K C S Q Q S E L [I A N	90		
Sc_Est2p	36	E I L T T C F A L P N S R - K I A L P C L P G D L S H K A V - D H	67		
EA_p123	62	I V A T P R D Y N E E D F K V I A R K E V F S T G L M I E L I D K	94		
Sp_Tip1p	91	V V K Q M F D E S F E R R - N L L M K [G F S M N H E D F R A M H	122		
Sc_Est2p	68	C I Y L L T G E L Y N - - N V L T F G Y K I A R N E D - - -	93		
EA_p123	95	C L V E L L S S S D V S D R Q K L Q C F G F Q L K G N Q - - -	122		
Sp_Tip1p	123	V N G V Q N D L V S T F P N Y L I S I L E S K N [W Q L L L E I I G	155		
Sc_Est2p	68	- - V N N S L F C H S A N V N V T L L K G A A W K M F H S L V G	123		
EA_p123	95	- - L A K T H L L T A L S T Q K Q Y F F Q D E W N Q V R A M I G	152		
Sp_Tip1p	123	S D A M H Y [L L S K G S I F E A L P N D N Y L Q I S G I I P L F K N	188		
Sc_Est2p	94	T Y A F V D L L I N Y T V I Q F N - G Q F F T Q I V G N R C N E P	155		
EA_p123	123	N E L F R H L Y T K Y L I F Q R T S E G T L V Q F C G N N V F D H	185		
Sp_Tip1p	156	N V F E E T V S K K R K R T I E T S I T Q N - - K S A R K E V S	218		
Sc_Est2p	124	H L P P K W V Q - - R S S S S A T A A Q I - - K Q L T E P V T	183		
EA_p123	153	L K V N D K F D K - - K Q K G G A A D M N E P R C C S T C K Y N V K	217		

FIG. 64
(CONTINUED)

B.

Sp_Tip1p	219	WN S I S I S R F S I F Y R S S Y K K F K Q D L Y F N L H S I C D
Sc_Est2p	184	N - - - - -
Ea_p123	218	N E K - - D H F L N N I N V P N W N N M K S R T R I F Y C T H F N
Sp_Tip1p	252	R N T V H M W L Q W I F P R Q F G L I N A F Q V K Q L H K V I P L
Sc_Est2p	201	- - - - - Y S K I L P S S - - S I K K L T D L R E A I F P
Ea_p123	249	R - - - - - N N Q F F K K H E F V S N K N N I S A M D R A Q T I
Sp_Tip1p	285	V S - - - Q S T V V P K R L L K V Y P L I E Q T A K R L H R I S
Sc_Est2p	224	T N - - - L V K I P Q R L K V R I N L T L Q K L L K R H K R L N
Ea_p123	276	F T N I F R F N R I R K K L K D K V I E K I A Y M L E K V K D F N
Sp_Tip1p	314	L S K V Y N H Y C P Y I D - T H D D E K I L S Y S L K P N Q - - -
Sc_Est2p	253	Y V S I L N S I C P P L E G T V L D L S H L S R Q S P K E R - - -
Ea_p123	309	F N Y Y L T K S C P L P E N W R E R K Q K I E N L I N K T R E E K
Sp_Tip1p	343	- - - - - - - - - - -
Sc_Est2p	283	- - - - - - - - - - -
Ea_p123	342	S K Y Y E E L F S Y T T D N K C V T Q E I N E F F Y N I L P K D F
Sp_Tip1p	360	W G N Q R I F E I I L K D L E T F L K L S R Y E S F S L H Y L M S
Sc_Est2p	300	F G S K K N K G K I I K N L N L L S L P L N G Y L P F D S L L K
Ea_p123	375	L T G - R N R K N F Q K K V K K Y V E L N K H E L I H K N L L L E
Sp_Tip1p	393	N I K I S E I E W L V L G K R S N A K M C L S D F E K R K Q I F A
Sc_Est2p	333	K L R L K D F R W L F I S - - D I W F T K H N F E N L N Q L A I
Ea_p123	407	K I N T R E I S W M Q Q V E T S - A K H F Y Y F D H E N - I Y V L W

FIG. 64
(CONTINUED)

B.		Sp_Tip1p	426	EF I YWL YNSF I P L Q SFF Y T E SSDL RNR T V Y	458
Sc_Est2p	363	CF I SWL FRQL PK QTFF YCTE SSTVT - I V Y	394		
Ea_p123	438	KLLRW I FEDLVVSL RCFFYYVTEQQKSYSKTY	470		
Sp_Tip1p	459	FRKD I WKLLCRRPF TSMKMEAFAEKINENNVRMD	491		
Sc_Est2p	395	FRHDTWNKLITPFI VYEYFKTYLVENNVCRNHNNS	427		
Ea_p123	471	YRKNIWDVIMKMSIAADLKKETLAEVQEKEVEEW	503		
Sp_Tip1p	492	TQKTTLPAAVIRLLPKK NTFRL TNLRKRFLL	522		
Sc_Est2p	428	YTLSNFNHSKMR PPKSNNNEFR A PCRGAD	460		
Ea_p123	504	KKSLGFAPGKLRL PKK - TTFRP MTFNKKIV	534		
Sp_Tip1p	523	IKMGSNKKMLVSTNQTLRPPVASICHL INE - -	552		
Sc_Est2p	461	EEE - - FTIYKENHKNAIQPTQK LEYLRNKRPT	491		
Ea_p123	535	NSD - - RKTTKLTTNTKLLNSHLM KTLKNR - MF	564		
Sp_Tip1p	553	ESSGIPFNLLEVYMKLLTFKKDLL KHRMFGR - KK	584		
Sc_Est2p	492	SFTKIYSPTQIADR KEFKQRLL KKFNNVLPFL	524		
Ea_p123	565	KDPFGFAVFNYDDVMKKYEEFVCKWKQVGQPKL	597		
Sp_Tip1p	585	YFVRIDIKSCYDR KQDLMFR VKKKLKDPE - F	616		
Sc_Est2p	525	YFMKF DVKSCYDS PRMECMRI LKDALKNENG	557		
Ea_p123	598	FFATMD EKCYD SVNREKLSTFLKTTKLSSDF	630		
Sp_Tip1p	617	VIRKYATIHATSDRATKN - - - - -	634		
Sc_Est2p	558	FVRSSQYFFNTNTG - - - - -	570		
Ea_p123	631	WIMTAQILKRKNNIVIDSKNFRKKEMKDYFRQK	663		

FIG. 64
(CONTINUED)

+

76/103

B.			
Sp_Tip1p	635	F V S E A F S Y F D M V P F E K V V Q L L S - - M K T S D T L F V	665
Sc_Est2p	571	- - - - V L K L F N V V N A S R - - V P K P Y E L Y I	591
Ea_p123	664	F Q K I A L E G G Q Y P T L F S V L E Q N D L N A K K T L I V	696
Sp_Tip1p	666	D F V D Y W T K S S S E I F K M L K E H L S G H I V K I G N S Q Y	698
Sc_Est2p	592	D N V R T V H L S N Q D V I N V V E M E I F K T A L W V E D K C Y	624
Ea_p123	697	E A K Q R N Y F K K D N L L Q P V I N I C Q Y N Y I N F N G K F Y	729
Sp_Tip1p	699	L Q K V G I P Q G S I L S S F L C H F Y M E D L I D E Y L S F T K	731
Sc_Est2p	625	I R E D G L F Q G S S L S A P I V D L V Y D D L L E F Y S E F K A	657
Ea_p123	730	K Q T K G I P Q G L C V S S I L S S F Y Y A T L E E S S L G F L R	762
Sp_Tip1p	732	K K G - - - - S V L L R V V D D F L F I T V N K K D A K K	756
Sc_Est2p	658	S P S Q D - - - - T L I L K L A D D F L I I S T D Q Q Q V I N	684
Ea_p123	763	D E S M N P E N P N V N L L M R L T D D Y L L I T T Q E N N A V L	795
Sp_Tip1p	757	F L N I S L R G F E K H N F S T S L E K T V I N F E N S N G - - -	786
Sc_Est2p	685	I K K L A M G G F Q K Y N A K A N R D K I L A V S S Q Q S D - - -	713
Ea_p123	796	F I E K L I N V S R E N G F K F N M K K L Q T S F P L S P S K F A	828
Sp_Tip1p	787	- - - I I N N T F F N E S K K R M P F F G F S V N M R S L D T L L	816
Sc_Est2p	714	- - - D D T V I Q F C A - - M H I F V K E L E V W K H S S T M	739
Ea_p123	829	K Y G M D S V E E Q N I V Q D Y C D W I G I S I D M K T L A L M P	861
Sp_Tip1p	817	A C P K I D E A L F N S T S V E L T K H M G K S F F Y K I L R S S	849
Sc_Est2p	740	N N F H I R S K S S K G I F R S L I A L F N T R I S Y K T I D T N	772
Ea_p123	862	N I N L R I E G I L C T L N L N M Q T K K A S M W L K K K L K S F	894

FIG. 64
(CONTINUED)

+

B.							
Sp_Tip1p	850	[L]A S F A Q V F I D I T H N S K F N S C C N I Y R L G Y S M C M R	882				
Sc_Est2p	773	[L]N S T N T V L M Q I D H V V K N I S E C - - - - -	793				
Ea_p123	895	[L]M N N I T H Y F R K T I T T E D F A N K T L N K L F I S G G Y K	927				
Sp_Tip1p	883	A Q A Y L K R M [K]D I F I P Q R M F I T D L L N V I G R K [I]W K K	915				
Sc_Est2p	794	- - - Y K S A F K D L S I N - V T Q N M Q F H S F L Q R I I E M	821				
Ea_p123	928	Y M Q C A K E Y K D H F K K N L A M S S M I D L E V S K I I Y S V	960				
Sp_Tip1p	916	L A E I L G Y T S R R F L S S A E V K W L F C L G M R D G L [K]P S	948				
Sc_Est2p	822	T V S G C P I T K C D P L I E Y E V R F T I L N G F L E S L S S N	854				
Ea_p123	961	T R A F F K Y L V C N I K D T I F G E E H Y P D F F L S T [L]K H F	993				
Sp_Tip1p	949	F K Y H P C F E Q L I Y Q [F]Q S L T D L I K P L R P V [L]R Q V L F	981				
Sc_Est2p	855	T S - - - - - K F K D N I I L L R K E I Q H L Q A Y I Y	877				
Ea_p123	994	I E I F S - - - T K K Y I F N R V C M I L K A K E A K [L]K S D Q C	1023				
Sp_Tip1p	982	L H R R I A D -	988				
Sc_Est2p	878	I Y I H I V N -	884				
Ea_p123	1024	Q S L I Q Y D A	1031				

FIG. 64
(CONTINUED)

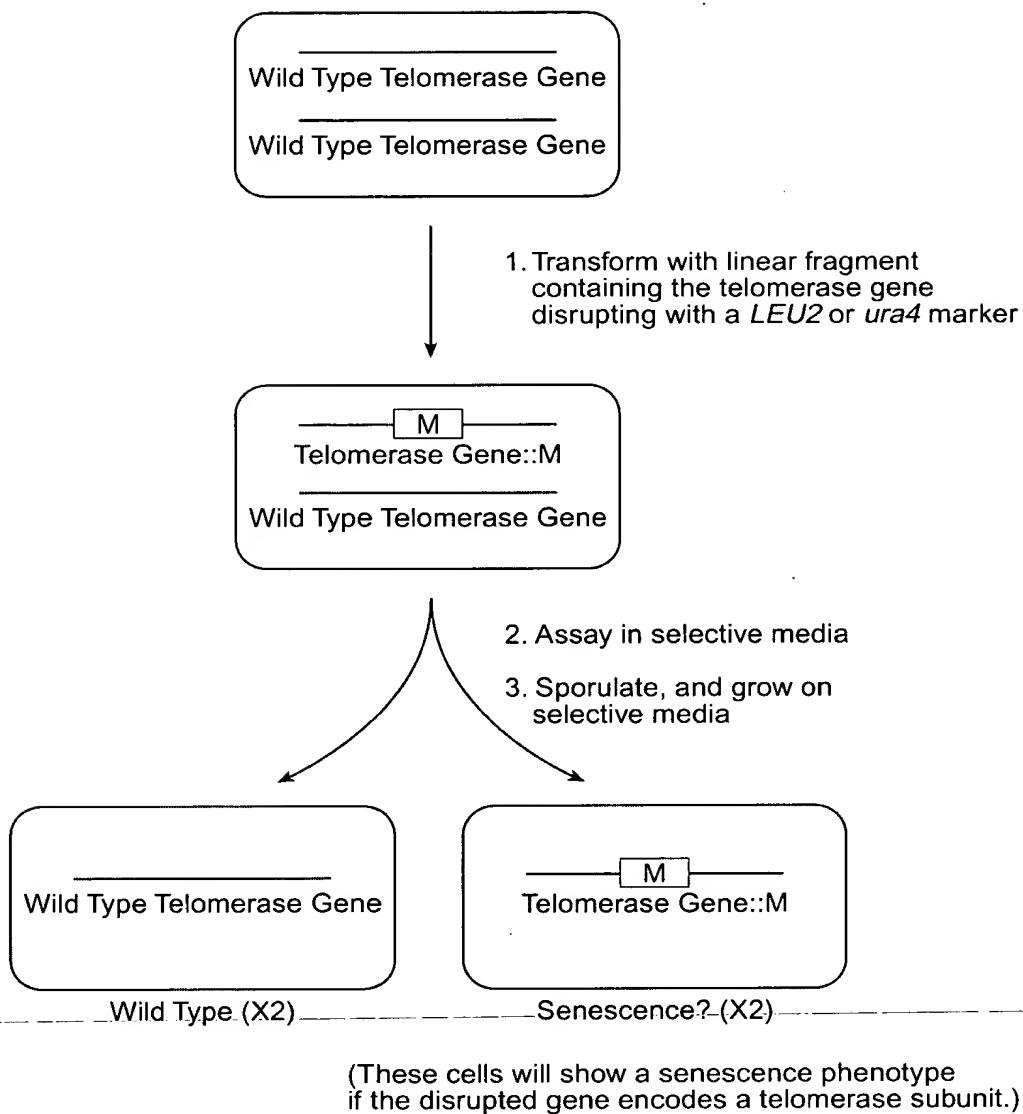


FIG. 65

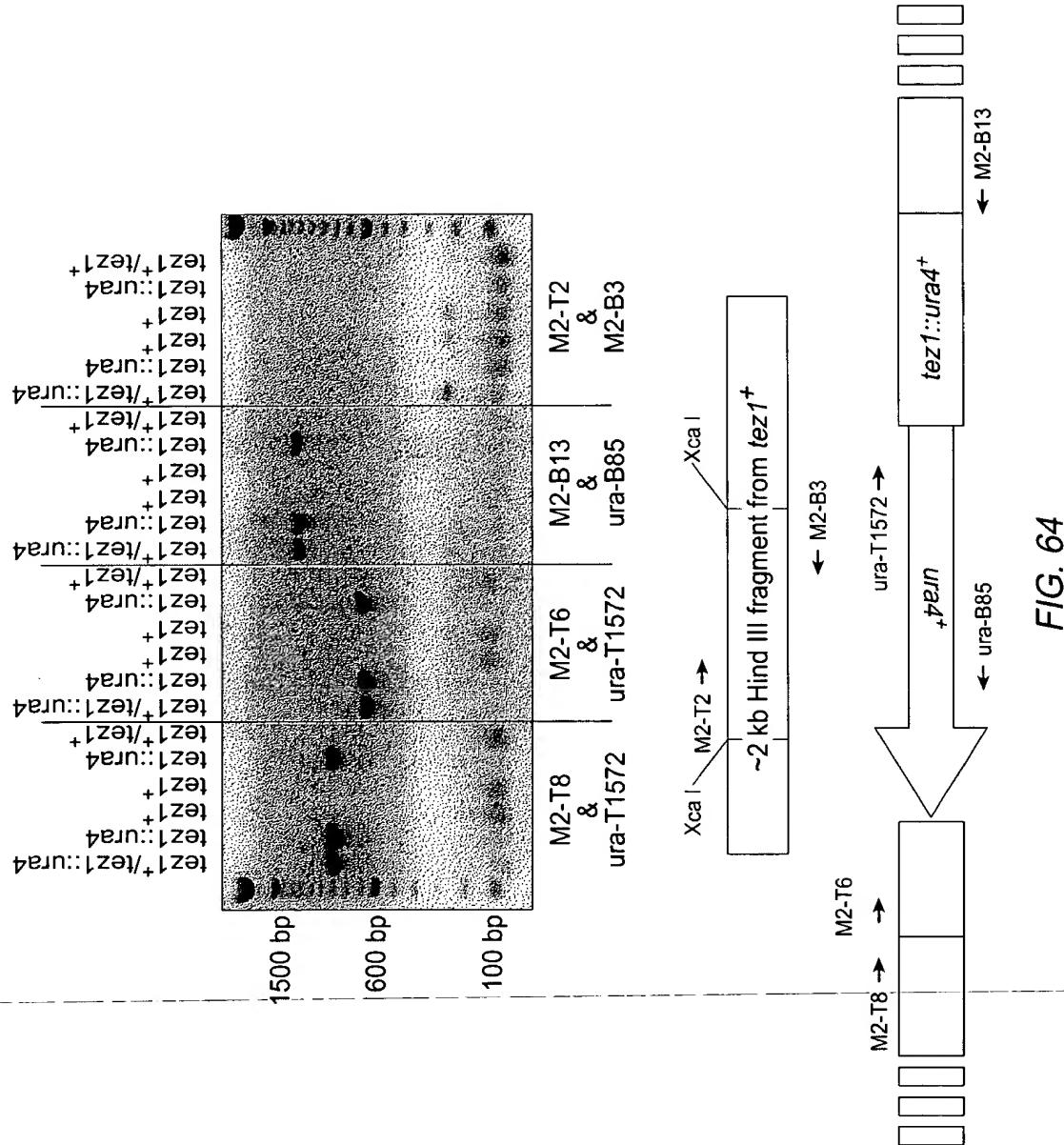


FIG. 64

+

80/103

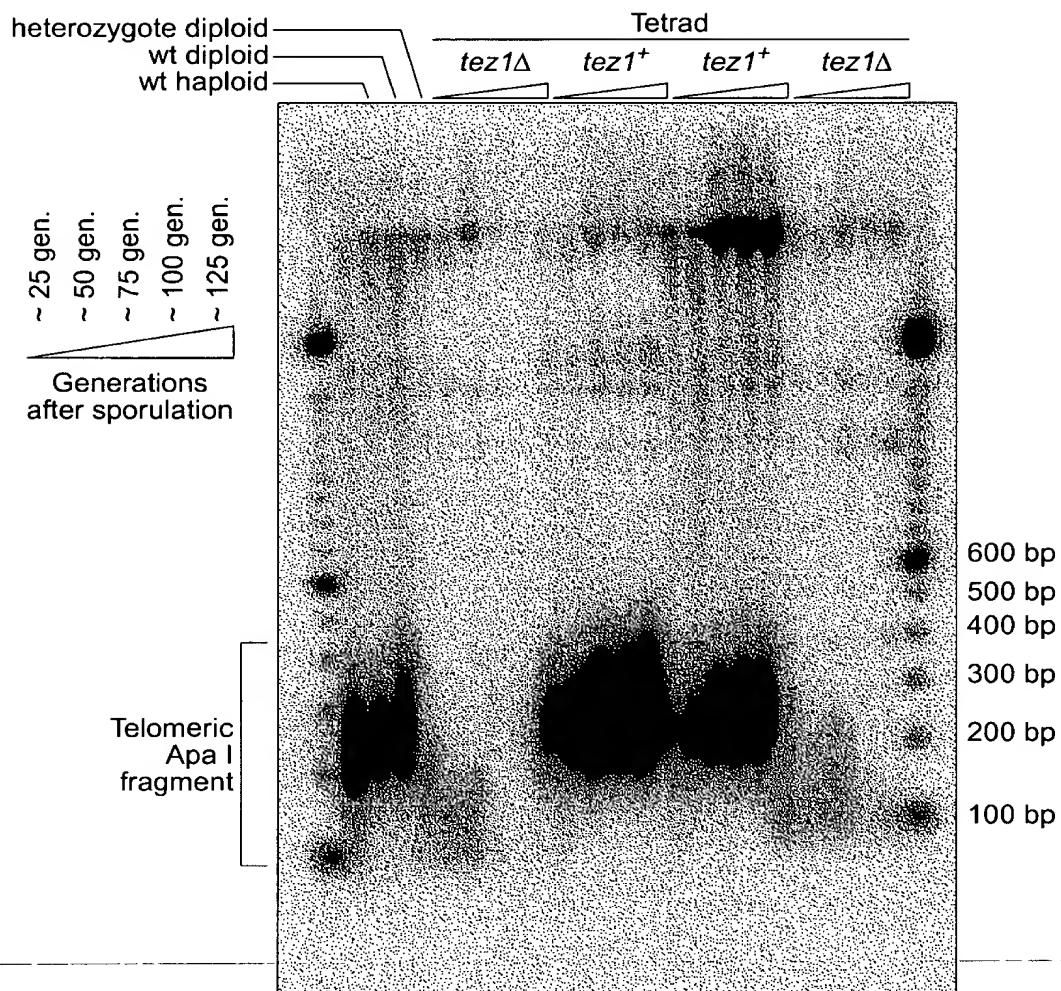


FIG. 67

+

1
GCCAAGTTCCCTGCACTGGCTG met ser val tyr val val glu leu leu
ATG AGT GTG TAC GTC GTC GAG CTG CTC

10
arg ser phe phe tyr val thr glu thr thr phe gln lys asn arg
AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG

20
leu phe phe tyr arg lys ser val trp ser lys leu gln ser ile
CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT

30
gly ile arg gln his leu lys arg val gln leu arg glu leu ser
GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG TCG

40
glu ala glu val arg gln his arg glu ala arg pro ala leu leu
GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG CTG

50
thr ser arg leu arg phe ile pro lys pro asp gly leu arg pro
ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG CCG

60
ile val asn met asp tyr val val gly ala arg thr phe arg arg
ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC AGA

70
glu lys ala glu arg leu thr ser arg val lys ala leu phe
GAA AAG ARG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG TTC

80
ser val leu asn tyr glu arg ala arg arg pro gly leu leu gly
AGC GTG CTC AAC TAC GAG CGG GCG CGC CCC GGC CTC CTG GGC

90
ala ser val leu gly leu asp asp ile his arg ala trp arg thr
GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC ACC

100
phe val leu arg val arg ala gln asp pro pro pro glu leu tyr
TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG TAC

110
phe val lys val asp val thr gly ala tyr asp thr ile pro gln
TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG

120
asp arg leu thr glu val ile ala ser ile ile lys pro gln asn
GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC AAA CCC CAG AAC

130
thr tyr cys val arg arg tyr ala val val gln lys ala ala met
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG

140

150
160 _____ 170 _____
phe val lys val asp val thr gly ala tyr asp thr ile pro gln
TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG

180
asp arg leu thr glu val ile ala ser ile ile lys pro gln asn
GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC AAA CCC CAG AAC

190
200
thr tyr cys val arg arg tyr ala val val gln lys ala ala met
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG

FIG. 68

+

82/103

210

gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys
 GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC

220

gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser
 CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC

240

leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg
 CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG

250

arg asp gly leu leu leu arg leu val asp asp phe leu leu val
 CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT TTC TTG GTG

270

thr pro his leu thr his ala lys thr phe leu arg thr leu val
 ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC

280

arg gly val pro glu tyr gly cys val val asn leu arg lys thr
 CGA GGT GTC CCT GAG TAT GGC TGC GTG AAC TTG CGG AAG ACA

300

val val asn phe pro val glu asp glu ala leu gly gly thr ala
 GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT

310

phe val gln met pro ala his gly leu phe pro trp cys gly leu
 TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG

330

leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser
 CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC

340

tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly
 TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC

360

phe lys ala gly arg asn met arg arg lys leu phe gly val leu
 TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG

370

arg leu lys cys his ser leu phe leu asp leu gln val asn ser
 CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC

380

leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln
 CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

400

ala tyr arg phe his ala cys val leu gln leu pro phe his gln
 GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

410

FIG. 68
(CONTINUED)

+

+

83/103

420

gln val trp lys asn pro his phe ser cys ala ser ser leu thr
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430

arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

450

val ala gly gly gln gly arg arg arg pro ser ala leu arg gly
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460

arg ala val ala val pro pro ser ile pro ala gln ala asp ser
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

480

thr pro cys his leu arg ala thr pro gly val thr gln asp ser
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490

pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GCT GAC TGC

510

pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520

his pro gly leu met ala thr arg pro gln pro gly arg glu gln
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

540

thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550

arg gly gly pro his pro gly leu his arg trp glu ser glu ala
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

564

OP

TGA GTGAGTGTGTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC

CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTCACTTCCCCAC

AGGCTGGCGTTGGTCCACCCCAGGGCCAGCTTCTCACCAAGGAGCCGGCTTCCACT

CCCCACATAGGAATAGTCCATCCCCAGATTGCCATTGTTCACCCCTGCCCTGCCCTCC

TTTGCCTTCCACCCCCACCATTCAAGGTGGAGACCCCTGAGAAGGACCTGGAGCTTGAG

AATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCCTGCACCTGGATGGGG

GTCCCTGTGGGTCAAATTGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATGAGTT

TTTCAGTTTGGAAAAAAAAAAAAAA

FIG. 68
(CONTINUED)

+

Motif -1	
Ep p123	...LVVSLIRCFYVTEQQKSYSKT...
Sp Tez1	...FIIPILQSFYITESSDLRNRT...
Sc Est2	...LIPKIIQTFFYCTEISSTVTIV...
Hs TCP1	...YVVELLRSFYVTETTFQKNRL...
consensus	FFY TE
Motif 0	K
Ep p123	p hhh K hR h R
Sp Tez1	...KSLGFAPGKLRLIPKKT--TFRPIMTFNKKIV...
Sc Est2	...QKTTLPPAVIRLLPKKN--TFRLLITNLRKRF...
Hs TCP1	...TLSNFNHSKMRRIIPKKSNNEFRIIAIAPCRGAD...
consensus	...ARPALLTSRLRFIPKPD--GLRPIVNMDYVVG...
	R PK R I
Motif A	AF
Ep p123	h hDh GY h
Sp Tez1	...PKLFFATMDIEKYDSVNREKLSTFLK...
Sc Est2	...RKKYFVRIDIKSCYDRIKQDLMFRIVK...
Hs TCP1	...PELYFMKFDVKSCYDSIPRMECMRILK...
consensus	...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...
	F D YD
Motif B	hPQG ps hh
Ep p123	...NGKFYKQTKGIPQGLCVSSILSSFYA...
Sp Tez1	...GNSQYLQKVGIHQGSILSSFLCHFYME...
Sc Est2	...EDKCYIREDGLFQGSSLsapIVDLVYD...
Hs TCP1	...RATSYVQCQGIPQGSILSTLCSLCYG...
consensus	G QG S
Motif C	Y
Ep p123	h F DD hhh
Sp Tez1	...PNVNLLMRLTDYLLITTQENN...
Sc Est2	...KKGSVLLRVVDDFLFITVNKKD...
Hs TCP1	...SQDTLILKLAADDFLIISTDQQQ...
consensus	...RRDGLLLRLVDDFLLVTPHLTH...
	DD L
Motif D	Gh h cK
Ep p123	...NVSRENGFKFNMKKL...
Sp Tez1	...LNLSLRGFEKHNFST...
Sc Est2	...KKLAMGGFQKYNKA...
Hs TCP1	...LRTLVRGVPEYGCVV...
consensus	G

FIG. 69

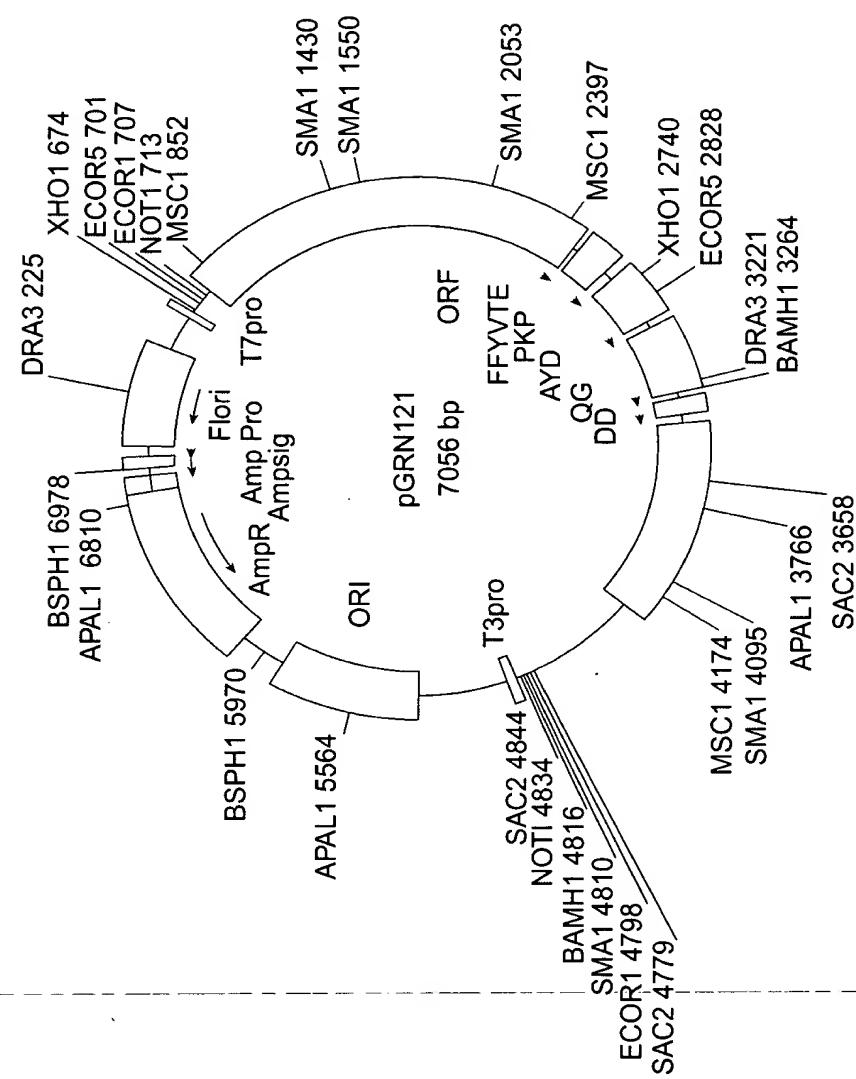


FIG. 70

1	GCAGCGCTGC	GTCCTGCTGC	GCACGTGGGA	AGCCCTGGCC	CCGGCCACCC
51	CCGCGATGCC	GCGCGCTCCC	CGCTGCCGAG	CCGTGCGCTC	CCTGCTGCAC
101	AGCCACTACC	GCGAGGTGCT	GCCGCTGGCC	ACGTTCTGTC	GGCGCCTGGG
151	GCCCCAGGGC	TGGCGGCTGG	TGCAGCGCG	GGACCCGGCG	GCTTCCGCG
201	CGNTGGTGGC	CCANTGCNTG	GTGTGCGTGC	CCTGGGANGN	ANGGCNGCCC
251	CCCGCCGCC	CCTCCCTCCG	CCAGGTGTCC	TGCCTGAANG	ANCTGGTGGC
301	CCGAGTGTG	CANANGCTGT	GCGANCGCG	CGCGAANAAC	GTGCTGGCCT
351	TCGGCTTCGC	GCTGCTGGAC	GGGGCCCGCG	GGGGCCCCCCC	CGAGGCCCTTC
401	ACCACCAGCG	TGCGCAGCTA	CCTGCCAAC	ACGGTGACCG	ACGCACTGCG
451	GGGGAGCGGG	GCGTGGGGGC	TGCTGCTGCG	CCGCGTGGC	GACGACGTGC
501	TGGTTCACCT	GCTGGCACGC	TGCGCGNTNT	TTGTGCTGGT	GGNTCCCAGC
551	TGCGCCTACC	ANGTGTGCGG	GCCGCCGCTG	TACCAGCTCG	GCGCTGCNAC
601	TCAGGCCCGG	CCCCCGCCAC	ACGCTANTGG	ACCCGAANGC	GTCTGGGATC
651	CAACGGGCCT	GGAACCATAG	CGTCAGGGAG	GCCGGGGTCC	CCCTGGGCTG
701	CCAGCCCCGG	GTGCGAGGAG	GCGCGGGGGC	AGTGCCAGCC	GAAGTCTGCC
751	GTTGCCAAG	AGGCCAGGC	GTGGCGCTGC	CCCTGAGCCG	GAGCGGACGC
801	CCGTTGGGCA	GGGGTCCCTGG	GCCCACCCGG	GCAGGACGCC	TGGACCGAGT
851	GACCGTGGTT	TCTGTGTGGT	GTCACCTGCC	AGACCCGCCG	AAGAAGCCAC
901	CTCTTGGAG	GGTGCCTCT	CTGGCACGCC	CCACTCCCAC	CCATCCGTGG
951	GCCGCCAGCA	CCACGGGGGC	CCCCCATCCA	CATCCGGGCC	ACCACGTCC
1001	GGGACACGCC	TTGTCCCCCG	GTGTACCGCG	AGACCAAGCA	CTTCCTCTAC
1051	TCCTCAGGCC	ACAAGNACAC	TGCGNCCCTC	CTTCCTACTC	AATATATCTG
1101	AGGCCCAGCC	TGACTGGCGT	TCGGGAGGTT	CGTGGAGACA	NTCTTCTGG
1151	TTCCAGGCC	TGGATGCCAG	GATTCCCCGC	AGTTGCCCC	GCCTGCCCA
1201	GCGNTACTGG	CAAATGCGC	CCCTGTTCT	GGAGCTGCTT	GGGAACCACG
1251	CGCAGTGCC	CTACGGGGTG	TTCCCTAAAGA	CGCACTGCC	GCTGCGAGCT
1301	GCGGTACACC	CAGCAGCCGG	TGTCTGTGCC	CGGGAGAAGC	CCCAGGGCTC
1351	TGTGGCGGCC	CCCGGAGGAG	AGGAACACAG	ACCCCGTCG	CCTGGTGCAG
1401	CTGCTCCGCC	AGCACAGCAG	CCCCTGGCAG	GTGTACGGCT	TCGTGCGGGC
1451	CTGCCTGCC	CGGCTGGTGC	CCCCAGGCCT	CTGGGGCTCC	AGGCACAACG
1501	AACGCCGCTT	CCTCAGGAAC	ACCAAGAAGT	TCATCTCCCT	GGGAAAGCAT
1551	GCCAAGCTCT	CGCTGCAGGA	GCTGACGTGG	AAGATGAGCG	TGCGGGACTG
1601	CGCTTGGCTG	CGCAGGAGCC	CAGGGGTTGG	CTGTGTTCCG	GCCGCAGAGC
1651	ACCGTCTGCC	TGAGGAGATC	CTGGCCAAGT	TCTGCACTG	GCTGATGAGT
1701	GTGTACGTG	TCGAGCTGCT	CAGGTCTTTC	TTTTATGTCA	CGGAGACCCAC
1751	GTTCAAAAG	AACAGGCTCT	TTTTCTACCG	GAAGAGTGTG	TGGAGCAAGT
1801	TGCAAAGCAT	TGGAATCAGA	CAGCACTGA	AGAGGGTGC	GCTGCGGGAG
1851	CTGTCGGAAG	CAGAGGTCA	GCAGCATCGG	GAAGCCAGGC	CCGCCCTGCT
1901	GACGTCCAGA	CTCCGCTTCA	TCCCCAAGCC	TGACGGGCTG	CGGCCGATTG
1951	TGAACATGG	CTACGTCGTG	GGAGCCAGAA	CGTTCGCGAG	AGAAAAGAGG
2001	GCCGAGCGTC	TCACCTCGAG	GGTGAAGGCA	CTGTTCAGCG	TGCTCAACTA
2051	CGAGCGGGCG	CGGCGCCCCG	GCCTCCTGGG	CGCCTCTGTG	CTGGGCCCTGG
2101	ACGATATCCA	CAGGGCCTGG	CGCACCTTCG	TGCTGCGTGT	GCAGGGCCCA
2151	GACCCGCCGC	CTGAGCTGTA	CTTGTCAAG	GTGGATGTGA	CGGGCGCGTA
2201	CGACACCATC	CCCCCAGGACA	GGCTCACCGA	GGTCATCGCC	AGCATCATCA
2251	AACCCCAGAA	CACGTACTGC	GTGCGTCGGT	ATGCCGTGGT	CCAGAAGGCC
2301	GCCCATGGGC	ACGTCCGAA	GGCCTCAAG	AGCCACGTCT	CTACCTTGAC
2351	AGACCTCCAG	CCGTACATGC	GACAGTCGT	GGCTCACCTG	CAGGANAACA
2401	GCCCGCTGAG	GGATGCCGTC	GTCATCGAGC	AGAGCTCCTC	CCTGAATGAG
2451	GCCAGCAGTG	GCCTCTCGA	CGTCTTCCTA	CGCTTCATGT	GCCACCAACGC

FIG. 71

2501 CGTGCATC AGGGCAAGT CCTACGCTTA GTGCCAGGGG ATCCCGCAGG
 2551 GCTCCATCCT CTCCACGCTG CTCTGCAGCC TGTGCTACGG CGACATGGAG
 2601 AACAAAGCTGT TTGCGGGGAT TCGGCAGGAC GGGCTGCTCC TGCGTTGGT
 2651 GGATGATTC TTGTTGGTGA CACCTCACCT CACCCACGCG AAAACCTTCC
 2701 TCAGGACCTT GGTCCGAGGT GTCCCTGAGT ATGGCTGCGT GGTGAACCTG
 2751 CGGAAGACAG TGGTGAACCT CCCTGTAGAA GACGAGGCC TGGGTGGCAC
 2801 GGCTTTGTT CAGATGCCGG CCCACGGCT ATTCCCTGG TGCGGCCTGC
 2851 TGCTGGATAC CCGGACCCCTG GAGGTGCAGA GCGACTACTC CAGCTATGCC
 2901 CGGACCTCCA TCAGAGCCAG TCTCACCTTC AACCAGGGCT TCAAGGCTGG
 2951 GAGGAACATG CGTCGCAAAC TCTTGGGGT CTTGCGGCTG AAGTGTACA
 3001 GCCTGTTCT GGATTGCAAG GTGAACAGCC TCCAGACGGT GTGCACCAAC
 3051 ATCTACAAGA TCCTCCTGCT GCAGGCGTAC AGGTTTACAG CATGTGTGCT
 3101 GCAGCTCCCCA TTTCATCAGC AAGTTGGAA GAACCCCACA TTTTCCTGC
 3151 GCGTCATCTC TGACACGGCC TCCCTCTGCT ACTCCATCCT GAAAGCCAAG
 3201 AACGCAGGGA TGTGCGCTGGG GGCCAAGGGC GCCGCCGGCC CTCTGCCCTC
 3251 CGAGGCCGTG CAGTGGCTGT GCCACCAAGC ATTCTGCTC AAGCTGACTC
 3301 GACACCGTGT CACCTACGTG CCACTCTGG GGTCACTCAG GACAGCCCAG
 3351 ACGCAGCTGA GTCGGAAGCT CCGGGGGACG ACGCTGACTG CCCTGGAGGC
 3401 CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT
 3451 GGCACCCCGC CCACAGCCAG GCGAGAGCA GACACCAGCA GCCCTGTAC
 3501 GCCGGGCTCT ACGTCCCAGG GAGGGAGGGG CGGCCACAC CCAGGCCCCGC
 3551 ACCGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCGAG GCCTGCATGT
 3601 CCGGCTGAAG GCTGAGTGTG CCGCTGAGGC CTGAGCGAGT GTCCAGCCAA
 3651 GGGCTGAGTG TCCAGCACAC CTGCCGTCTT CACTTCCCCA CAGGCTGGCG
 3701 CTCGGCTCCA CCCCCAGGGCC AGCTTTCTT CACCAAGGAGC CCGGCTTCCA
 3751 CTCCCCACAT AGGAATAGTC CATCCCCAGA TTGCCCCATTG TTCACCCCTC
 3801 GCCTGCCCC CTTTGCCTT CCACCCCCAC CATCCAGGTG GAGACCCCTGA
 3851 GAAGGACCCCT GGGAGCTCTG GGAATTGGA GTGACCAAAG GTGTGCCCTG
 3901 TACACAGGCG AGGACCCCTGC ACCTGGATGG GGGTCCCTGT GGGTCAAATT
 3951 GGGGGGAGGT GCTGTGGGAG TAAAATACTG AATATATGAG TTTTCAGTT
 4001 TTGAAAAAAA AAAAAAAA AAAAAAAA

FIG. 71

(CONTINUED)

GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGCCACCCCCCGCGATGCC
 1 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
 CGTCGCGACGCAGGACGACGCGTGACCCCTCAGGGACCGGGCCGGTGGGGCGCTACGG

a A A L R P A A H V G S P G P G H P R D A -
 b Q R C V L L R T W E A L A P A T P A M P -
 c S A A S C C A R G K P W P R P P P R C R -

GCGCGCTCCCCGCTGCCGAGCCGTGCGCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCT
 61 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
 CGCGCGAGGGGCGACGGCTCGGCACCGCGAGGGACGACGCGTCGGTATGGCGCTCACGA

a A R S P L P S R A L P A A Q P L P R G A -
 b R A P R C R A V R S L L R S H Y R E V L -
 c A L P A A E P C A P C C A A T T A R C C -

GCCGCTGGCCACGTTCGTGCAGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGG
 121 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
 CGCGACCGGTGCAAGCACGCCCGACCCGGGTCCCACCGCCGACCACGTCGCGCC

a A A G H V R A A P G A P G L A A G A A R -
 b P L A T F V R R L G P Q G W R L V Q R G -
 c R W P R S C G A W G P R A G G G W C S A G -

GGACCCGGCGGCTTCCGCGCGNTGGTGGCCCANTGCNTGGTGTGCGTGCCTGGANGN
 181 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
 CCTGGGCCGCGAAAGGCGCGCNACCACCGGGTNACGNACCACACGACGGGACCCCTNCN

a G P G G F P R ? G G P ? ? G V R A L G ? -
 b D P A A F R A ? V A ? C ? V C V P W ? ? -
 c T R R L S A R W W P ? A W C A C P G ? ? -

ANGGCNGCCCCCGCCGCCCCCTCCTCCGCCAGGTGTCCCTGCCTGAANGANCTGGTGGC
 241 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
 TNCCGNCGGGGGGCGCGGGGGAGGAAGGCGGTCCACAGGACGGACTTNCTNGACCACCG

a ? A A P R R P L L P P G V L P E ? ? G G -
 b ? ? P P A A P S F R Q V S C L ? ? L V A -
 c G ? P P P P P S A R C P A * ? ? W W P -

CCGAGTGCTGCANANGCTGTGCGANCAGCGCGCGAANAACGTGCTGGCCTTCGGCTTCGC
 301 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 360
 GGCTCACGACGTNTNCGACACGCTNGCGCCCGCTTNTGCACGACCGGAAGCCGAAGCG

a P S A A ? A V R ? R R E ? R A G L R L R -
 b R V L ? ? L C ? R G A ? N V L A F G F A -
 c E C C ? ? C A ? A A R ? T C W P S A S R -

GCTGCTGGACGGGCCCGCGGGGGCCCCGAGGGCTTCAACCACGCGTGCGCAGCTA
 361 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 420
 CGACGACCTGCCCGGGCGCCCCGGGGGCTCCGGAAAGTGGTGGTGCACGCGTCGAT

a A A G R G P R G P P R G L H H Q R A Q L -
 b L L D G A R G G P P E A F T T S V R S Y -
 c C W T G P A G A P P R P S P P A C A A T -

CCTGCCAACACGGTGACCGACGCACTGCGGGGGAGCGGGGGCGTGGGGCTGCTGCG
 421 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 480
 GGACGGGTTGTGCCACTGGCTGCGTGACGCCCTCGCCCCGACCCCCGACGACGACGC

a P A Q H G D R R T A G E R G V G A A A A -
 b L P N T V T D A L R G S G A W G L L L R -
 c C P T R * P T H C G G A G R G G C C C A -

FIG. 72

a P R G R R R A G S P A G T L R ? ? C A G -
 b R V G D D V L V H L L A R C A ? F V L V -
 c A W A T T C W F T C W H A A R ? L C W W -

 541 GGNNTCCCAGCTGCGCCTACCANGTGTGCGGGCCGCCGCTGTACCAAGCTCGGCCTGCNAC
 CCNAGGGTCACGCGGATGGTNACACGCCGGCGACATGGTCAGGCCGACGNTG 600

 a G S Q L R L P ? V R A A A V P A R R C ? -
 b ? P S C A Y ? V C G P P L Y Q L G A A T -
 c ? P A A P T ? C A G R R C T S S A L ? L -

 601 TCAGGCCCGGCCCCGCCAACACGCTANTGGACCCGAANGCGTCTGGATCCAACGGGCCT
 AGTCCGGGCCGGGGCGGTGTGCGATNACCTGGCTTNCAGACCCTAGGTTGCCCGA 660

 a S G P A P A T R ? W T R ? R L G S N G P -
 b Q A R P P P H A ? G P E ? V W D P T G L -
 c R P G P R H T L ? D P ? A S G I Q R A W -

 661 GGAACCATAGCGTCAGGGAGGCCGGGTCCCCCTGGCTGCCAGCCCCGGTGCGAGGAG
 CCTTGGTATCGCAGTCCTCCGGCCCCAGGGGACCCGACGGTCGGGGCACGCTCCTC 720

 a G T I A S G R P G S P W A A S P G C E E -
 b E P * R Q G G R G P P G L P A P G A R R -
 c N H S V R E A G V P L G C Q P R V R G G -

 721 GCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCCAAGAGGCCAGGCCTGGCGCTGC
 CGCGCCCCCGTACCGTCGGCTTCAGACGGCAACGGGTTCTCCGGTCCGACCGCGACG 780

 a A R G Q C Q P K S A V A Q E A Q A W R C -
 b R G G S A S R S L P L P K R P R R G A A -
 c A G A V P A E V C R C P R G P G V A L P -

 781 CCCTGAGCCGGAGCGGACGCCGTTGGCAGGGTCTGGCCACCCGGCAGGACGCC
 GGGACTCGGCCTCGCCTGCCGGCAACCCGTCCCCAGGACCCGGTGGGCCGCTCGGG 840

 a P * A G A D A R W A G V L G P P G Q D A -
 b P E P E R T P V G Q G S W A H P G R T P -
 c L S R S G R P L G R G P G P T R A G R L -

 841 TGGACCGAGTGACCGTGGTTCTGTGTGGTGTACCTGCCAGACCCGCCGAAGAAC
 ACCTGGCTCACTGGCACCAAAGACACACCACAGTGGACGGTCTGGCGCTTCTCGGT 900

 a W T E * P W F L C G V T C Q T R R R S H -
 b G P S D R G F C V V S P A R P A E E A T -
 c D R V T V V S V W C H L P D P P K K P P -

 901 CTCTTGGAGGGTGCCTCTGGCACCGGCCACTCCACCCATCCGTGGCCAGCA
 GAGAAACCTCCACCGAGAGACCGTGCAGGGTAGGCACCCGGCGGT 960

 a L F G G C A L W H A P L P P I R G P P A -
 b S L E G A L S G T R H S H P S V G R Q H -
 c L W R V R S L A R A T P T H P W A A S T -

 961 CCACGCCGGCCCCCATCCACATCGCGGCCACCACGTCTGGACACGCCCTGTCCCC
 GGTGCGCCGGGGTAGGTGTAGCGCCGGTGGTCAGGACCCGTGCGGAACAGGGGC 1020

FIG. 72
(CONTINUED)

+

90/103

a P R G P P I H I A A T T S W D T P C P P -
 b H A G P P S T S R P P R P G T R L V P R -
 c T R A P H P H R G H H V L G H A L S P G -

1021 GTGTACGCCGAGACCAAGCAGTCTACTCCTCAGGCAGAAGNACACTGCGNCCTC
 CACATGCGGCTCTGGTCGTGAAGGAGATGAGGAGTCCGCTGTCNTGTGACGCNGGGAG 1080

a V Y A E T K H F L Y S S G D K ? T A ? L -
 b C T P R P S T S S T P Q A T ? T L R P S -
 c V R R D Q A L P L L L R R Q ? H C ? P P -

1081 CTTCTACTCAATATATCTGAGGCCAGCCTGACTGGCGTCGGAGGTTGTGGAGACA
 GAAGGATGAGTTATATAGACTCCGGTCGGACTGACCGCAAGCCCTCCAAGCACCTCTGT 1140

a L P T Q Y I * G P A * L A F G R F V E T -
 b F L L N I S E A Q P D W R S G G S W R ? -
 c S Y S I Y L R P S L T G V R E V R G D ? -

1141 NTCTTTCTGGTCCAGGCCTGGATGCCAGGATTCCCCGCAGGTTGCCCGCCTGCCCA
 NAGAAAGACCAAGGTCCGGAACCTACGGCTTAAGGGCGTCAAACGGGGCGACGGGT 1200

a ? F L V P G L G C Q D S P Q V A P P A P -
 b S F W F Q A L D A R I P R R L P R L P Q -
 c L S G S R P W M P G F P A G C P A C P S -

1201 GCGNTACTGGCAAATGCGGCCCTGTTCTGGAGCTGCTGGGAACCACGCGCAGTGCCC
 CGCNATGACCGTTACGCCGGGACAAGACCTCGACGAACCTTGGTGCACGGTACGGG 1260

a A ? L A N A A P V S G A A W E P R A V P -
 b R Y W Q M R P L F L E L L G N H A Q C P -
 c ? T G K C G P C F W S C L G T T R S A P -

1261 CTACGGGGTGTCTCAAGACGCACTGCCGCTGCGAGCTGCCGTACCCAGCAGCCGG
 GATGCCCAAGGAGTTCTGCGTGACGGCGACGCTCGACGCCAGTGGGTCTGGCC 1320

a L R G V P Q D A L P A A S C G H P S S R -
 b Y G V F L K T H C P L R A A V T P A A G -
 c T G C S S R R T A R C E L R S P Q Q P V -

1321 TGTCTGTGCCGGAGAACCCCCAGGGCTGTGGCGGCCCGAGGAGGAGGAACACAG
 ACAGACACGGCCCTCTCGGGGTCCCGAGACACCGCCGGGCTCCTCCTTGTGTC 1380

a C L C P G E A P G L C G G P R G G G T Q -
 b V C A R E K P Q G S V A A P E E E E H R -
 c S V P G R S P R A L W R P P R R N T D -

1381 ACCCCCCGTCGCTGGTGCAGCTGCTCCGCCAGCACAGCAGCCCCCTGGCAGGTGTACGGCT
 TGGGGCAGCGGACACGTCGACGAGGCGGTGTCGTCGTCGGGACCGTCCACATGCCGA 1440

a T P V A W C S C S A S T A A P G R C T A -
 b P P S P G A A A P P A Q Q P L A G V R L -
 c P R R L V Q L L R Q H S S P W Q V Y G F -

1441 TCGTGCAGGGCCTGCCTGCCGGCTGGTGCAGGCTCTGGGGCTCCAGGCACAACG
 AGCACGCCGGACGGACGCCGACCGGGGTCCGGAGACCCCGAGGGTCCGTGTTGC 1500

FIG. 72
 (CONTINUED)

+

+

91/103

FIG. 72
(CONTINUED)

1

+

92/103

a S P S L T G C G R L * . T W T T S W E P E -
b P Q A * R A A A D C E H G L R R G S Q N -
c P K P D G L R P I V N M D Y V V G A R T -

CGTTCCGAGAGAAAAGAGGGCCGAGCGTCAACCTCGAGGGTGAAGGCAGTGTCAGCG
1981 -----+-----+-----+-----+-----+-----+ 2040
GCAAGGGTCTCTTTCTCCGGCTCGCAGAGTGGAGCTCCACTTCCGTGACAAGTCGC

a R S A E K R G P S V S P R G * R H C S A -
b V P Q R K E G R A S H L E G E G T V Q R -
c F R R E K R A E R L T S R V K A L F S V -

TGCTCAACTACGAGCGGGCGCGCGCCCCGCTCCCTGGGCCCTCTGTGCTGGGCTGG
2041 -----+-----+-----+-----+-----+-----+ 2100
ACGAGTTGATGCTCGCCCGCGCCGCGGGGGCCGGAGGACCCGGAGACACGACCCGGACC

a C S T T S G R G A P A S W A P L C W A W -
b A Q L R A G A A P R P P G R L C A G P G -
c L N Y E R A R R P G L L G A S V L G L D -

ACGATATCCACAGGGCCTGGCGCACCTCGTGTGCGTGTGCGGGCCCAGGACCCGGCG
2101 -----+-----+-----+-----+-----+-----+ 2160
TGCTATAGGTGTCGGACCGCGTGGAAAGCACGACGCACAGCCGGTCTGGCGCG

a T I S T G P G A P S C C V C G P R T R R -
b R Y P Q G L A H L R A A C A G P G P A A -
c D I H R A W R T F V L R V R A Q D P P P -

CTGAGCTGTACTTGTCAAGGTGGATGTGACGGCGCGTACGACACCATCCCCCAGGACA
2161 -----+-----+-----+-----+-----+-----+ 2220
GACTCGACATGAAACAGTCCACCTACACTGCCCGCATGCTGTGGTAGGGGGTCTGT

a L S C T L S R W M * R A R T T P S P R T -
b * A V L C Q G G C D G R V R H H P P G Q -
c E L Y F V K V D V T G A Y D T I P Q D R -

GGCTCACGGAGGTATGCCAGCATCAAACCCAGAACACGTACTGCGTGCCTCGGT
2221 -----+-----+-----+-----+-----+-----+ 2280
CCGAGTGCCTCCAGTAGCGGCGTAGTAGTTGGGTCTGTGCATGACGCACGCCA

a G S R R S S P A S S N P R T R T A C V G -
b A H G G H R Q H H Q T P E H V L R A S V -
c L T E V I A S I I K P Q N T Y C V R R Y -

ATGCCGTGGTCCAGAACGGCGCCCATGGGACGTCCGCAAGGCCCTCAAGAGGCCACGTCT
2281 -----+-----+-----+-----+-----+-----+ 2340
TACGGCACCAAGGTCTCCGGCGGGTACCGTGCAGCGTCCCGAAGTTCTCGGTGCAGA

a M P W S R R P P M G T S A R P S R A T S -
b C R G P E G R P W A R P Q G L Q E P R L -
c A V V Q K A A H G H V R K A F K S H V S -

CTACCTTGACAGACCTCCAGCGTACATGCGACAGTTGTCGCTCACCTGCAGGANAACA
2341 -----+-----+-----+-----+-----+-----+ 2400
GATGGAACACTGTCGGAGGTGGCATGTACGCTGTCAAGCACCGAGTGGACGTCCCTNTTGT

a L P * Q T S S R T C D S S W L T C R ? T -
b Y L D R P P A V H A T V R G S P A G ? Q -
c T L T D L Q P Y M R Q F V A H L Q ? N S -

GCCCGCTGAGGGATGCCGTCGTACGAGCAGAGCTCCCTCCCTGAATGAGGCCAGCAGTG
2401 -----+-----+-----+-----+-----+-----+ 2460
CGGGCGACTCCCTACGGCAGTAGCTCGTCTCGAGGAGGGACTTACTCCGGTCGTAC

FIG. 72
(CONTINUED)

+

+

93/103

a A R * G M P S S S S R A P P * M R P A V -
b P A E G C R R H R A E L L P E * G Q Q W -
c P L R D A V V I E Q S S S L N E A S S G -

GCCTCTTCGACGTCTCCTACGTTCATGTGCCACCACGCCGTGCGCATCAGGGGCAAGT
2461 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2520
CGGAGAAGCTGAGAAGGATGCGAAGTACACGGTGGTGCGGCACGCGTAGTCCCCGTTCA

a A S S T S S Y A S C A T T P C A S G A S -
b P L R R L P T L H V P P R R A H Q G Q V -
c L F D V F L R F M C H H A V R I R G K S -

CCTACGTCCAGTGCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTGCAGCC
2521 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2580
GGATGCAGGTACGGTCCCTAGGGCGTCCCGAGGTAGGAGAGGTGCGACGAGACGTCGG

a P T S S A R G S R R A P S S P R C S A A -
b L R P V P G D P A G L H P L H A A L Q P -
c Y V Q C Q G I P Q G S I L S T L L C S L -

TGTGCTACGGCGACATGGAGAACAAAGCTGTTGCGGGGATTGGCGGGACGGGCTGCTCC
2581 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2640
ACACGATGCCGCTGTACCTCTGTTGACAAACGCCCTAACGCCCTGCCGACGAGG

a C A T A T W R T S C L R G F G G T G C S -
b V L R R H G E Q A V C G D S A G R A A P -
c C Y G D M E N K L F A G I R R D G L L L -

TGCGTTGGTGGATGATTCTTGTGTTGACACCTCACCTCACCCACGCGAAAACCTTCC
2641 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2700
ACGCAAACCACTACTAAAGAACAAACACTGTGGAGTGGAGTGGGTGCGCTTTGGAAGG

a C V W W M I S C W * H L T S P T R K P S -
b A F G G * F L V G D T S P H P R E N L P -
c R L V D D F L L V T P H L T H A K T F L -

TCAGGACCCCTGGTCCGAGGTGTCCTGAGTATGGCTGCGTGGTAACCTGCGGAAGACAG
2701 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2760
AGTCCTGGGACCAGGCTCACAGGGACTCATACCGACGCACCACTGAACGCCCTGTC

a S G P W S E V S L S M A A W * T C G R Q -
b Q D P G P R C P * V W L R G E L A E D S -
c R T L V R G V P E Y G C V V N L R K T V -

TGGTGAACCTCCCTGAGAACAGCAGGGCCCTGGTGGCACGGCTTTGTCAGATGCCGG
2761 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2820
ACCACTTGAAGGGACATCTCTGCTCCGGGACCCACCGTGCGCAAACAAGTCTACGGCC

a W * T S L * K T R P W V A R L L F R C R -
b G E L P C R R R G P G W H G F C S D A G -
c V N F P V E D E A L G G T A F V Q M P A -

CCCACGGCCTATTCCCTGGTGCAGGCGCTGCTGGATACCCGGACCCCTGGAGGTGCGAGA
2821 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2880
GGGTGCCGGATAAGGGGACCAACGCCGGACGACGACCTATGGCCTGGGACCTCCACGTCT

a P T A Y S P G A A C C C W I P G P W R C R -
b P R P I P L V R P A A G Y P D P G G A E -
c H G L F P W C G L L L D T R T L E V Q S -

GCGACTACTCCAGCTATGCCGGACCTCCATCAGAGCCAGTCTCACCTCAACCGGGCT
2881 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2940
CGCTGATGAGGTCGATAACGGGCTGGAGGTAGTCTCGGTCAAGAGTGGAAAGTTGGCGCCGA

FIG. 72
(CONTINUED)

+

1

FIG. 72
(CONTINUED)

a V G S S R G R R * L P W R P Q P T R H C -
 b S E A P G D D A D C P G G R S Q P G T A -
 c R K L P G T T L T A L E A A A N P A L P -

 CCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCACAGCCAGGCCAGAGCA
 3421 -----+-----+-----+-----+-----+-----+-----+ 3480
 GGAGTCTGAAGTTCTGGTAGGACCTGACTACCGGTGGGCCGGTGTGGTCCGGCTCTCGT

 a P Q T S R P S W T D G H P P T A R P R A -
 b L R L Q D H P G L M A T R P Q P G R E Q -
 c S D F K T I L D * W P P A H S Q A E S R -

 GACACCAGCAGCCCTGTACGCCGGCTCTACGTCCCAGGGAGGGAGGGCCGGCCACAC
 3481 -----+-----+-----+-----+-----+-----+-----+ 3540
 CTGTGGTCGTCGGACAGTGCGGCCGAGATGCAGGGTCCCTCCCTCCCCGCCGGGTGTG

 a D T S S P V T P G S T S Q G G R G G P H -
 b T P A A L S R R A L R P R E G G A A H T -
 c H Q Q P C H A G L Y V P G R E G R P T P -

 CCAGGCCCCGACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTGTTGGCCGAGGCCTGCATGT
 3541 -----+-----+-----+-----+-----+-----+-----+ 3600
 GGTCCGGCGTGGCGACCCCTCAGACTCCGACTCACTCACAAACCGGCTCCGGACGTACA

 a P G P H R W E S E A * V S V W P R P A C -
 b Q A R T A G S L R P E * V F G R G L H V -
 c R P A P L G V * G L S E C L A E A C M S -

 CCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTG
 3601 -----+-----+-----+-----+-----+-----+-----+ 3660
 GGCGACTTCCGACTCACAGGCCACTCCGACTCGCTCACAGGTGGTCCCGACTCAC

 a P A E G * V S G * G L S E C P A K G * V -
 b R L K A E C P A E A * A S V Q P R A E C -
 c G * R L S V R L R P E R V S S Q G L S V -

 TCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTGGCTCCACCCAGGGCC
 3661 -----+-----+-----+-----+-----+-----+-----+ 3720
 AGGTCGTGTGGACGGCAGAAGTGAAGGGGTGTCCGACCGCGAGCCGAGGTGGGTCCCGG

 a S S T P A V F T S P Q A G A R L H P R A -
 b P A H L P S S L P H R L A L G S T P G P -
 c Q H T C R L H F P T G W R S A P P Q G Q -

 AGCTTTTCTCACAGGAGCCCCGGCTTCACTCCCCACATAGGAATAGTCCATCCCCAGA
 3721 -----+-----+-----+-----+-----+-----+-----+ 3780
 TCGAAAAGGAGTGGTCTCGGGCCGAAGGTGAGGGTGTATCCTTATCAGGTAGGGTCT

 a S F S S P G A R L P L P T * E * S I P R -
 b A F P H Q E P G F H S P H R N S P S P D -
 c L F L T R S P A S T P H I G I V H P Q I -

 TTGCCCCATTGTTCACCCCTGCCCTGCCCTCCTTGCCTTCACCCACCATCCAGGTG
 3781 -----+-----+-----+-----+-----+-----+-----+ 3840
 AAGCGGTAACAAGTGGGAGCGGGACGGGAGGAACCGAAGGTGGGGGTGGTAGGTCCAC

 a F A I V H P S P C P P L P S T P T I Q V -
 b S P L F T P R P A L L C L P P P P S R W -
 c R H C S P L A L P S F A F H P H H P G G -

 GAGACCCCTGAGAAGGACCTGGGAGCTCTGGGAATTGGAGTGACCAAAGGTGTGCCCTG
 3841 -----+-----+-----+-----+-----+-----+-----+ 3900
 CTCTGGGACTCTTCCCTGGGACCCCTCGAGACCTAAACCTCACTGGTTCCACACGGGAC

FIG. 72
(CONTINUED)

a E T L R R T L G A L G I W S D Q R C A L -
 b R P * E G P W E L W E F G V T K G V P C -
 c D P E K D P G S S G N L E * P K V C P V -

TACACAGGCAGGGACCTGCACCTGGATGGGGTCCCTGTGGTCAAATTGGGGGAGGT
 3901 -----+-----+-----+-----+-----+-----+ 3960
 ATGTGTCCGCTCCTGGGACGTGGACCTACCCCCAGGGACACCCAGTTAACCCCCCTCCA

a Y T G E D P A P G W G S L W V K L G G G -
 b T Q A R T L H L D G G P C G S N W G E V -
 c H R R G P C T W M G V P V G Q I G G R C -

GCTGTGGGAGTAAAATACTGAATATATGAGTTTTCAGTTTGAAAAA 4020
 3961 -----+-----+-----+-----+-----+-----+ 4020
 CGACACCCTCATTTATGACTTATACTCAAAAAGTCAAAACCTTTTTTTTTTTTT

a A V G V K Y * I Y E F F S F E K K K K K -
 b L W E * N T E Y M S F S V L K K K K K K -
 c C G S K I L N I * V F Q F * K K K K K K K -

AAAAAAA
 4021 ----- 4029
 TTTTTTTT

a K K K -
 b K K -
 c K K -

FIG. 72
 (CONTINUED)

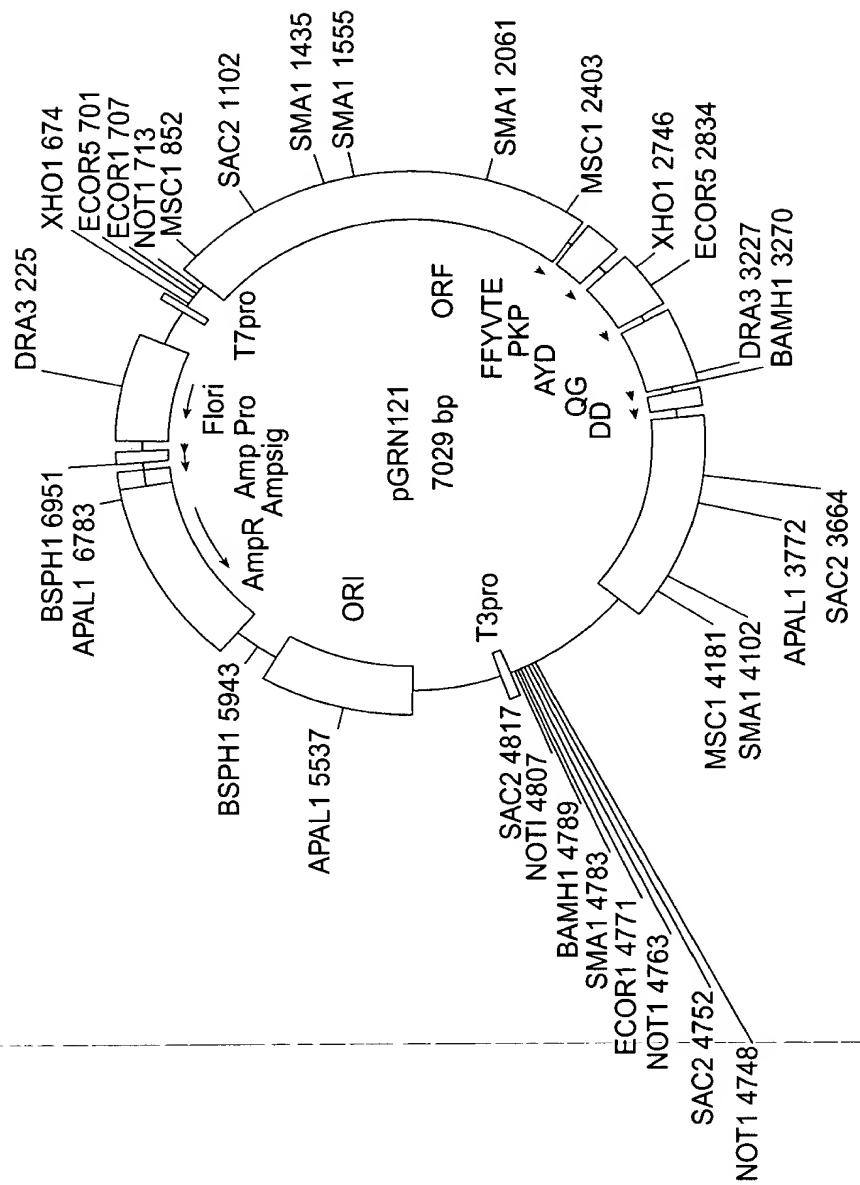


FIG. 73

+

98/103

1
met

GCAGCGCTGCGTCCCTGCTGCGCACGTGGAAAGCCCTGGCCCCGCCACCCCCCGCG ATG

10

pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

20

his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

30

gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

40

phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

50

ala arg pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

60

leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

70

gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG

80

ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC

90

tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

100

trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

110

ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC

120

tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

130

trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

140

ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC

150

leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

160

ala tyg gln val cys gly pro pro leu tyg gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

170

thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

180

190

FIG. 74

+

+

99/103

200

leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

220

val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

230

ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

250

ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

260

ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

280

val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

290

gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

310

gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

320

trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

330

340

leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

350

360

leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

370

glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

380

390

arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

400

leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

410

420

val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

FIG. 74
(CONTINUED)

+

+

100/103

430

ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

440

ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

460

leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

470

ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

490

his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

500

leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

520

met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

530

gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

550

ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

560

leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

580

arg leu phe phe tyr arg pro ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG CCG AGT GTC TGG AGC AAG TTG CAA AGC

590

ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

610

ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

620

leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

630

pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

640

FIG. 74
(CONTINUED)

+

101/103

650	660
arg glu lys arg ala glu arg leu thr ser arg val lys ala leu	
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG	
670	
phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu	
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CCC GGC CTC CTG	
680	
gly ala ser val leu gly leu asp asp ile his arg ala trp arg	690
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC	
700	
thr phe val leu arg val arg ala gln asp pro pro pro glu leu	
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG	
710	
tyr phe val lys val asp val thr gly ala tyr asp thr ile pro	720
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC	
730	
gln asp arg leu thr glu val ile ala ser ile ile lys pro gln	
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG	
740	
asn thr tyr cys val arg arg tyr ala val val gln lys ala ala	750
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC	
760	
his gly his val arg lys ala phe lys ser his val ser thr leu	
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC TCT ACC TTG	
770	
thr asp leu gin pro tyr met arg gln phe val ala his leu gln	780
ACA GAC CTC CAG CCG TAC ATG CGA CAG TTC GTG GCT CAC CTG CAG	
790	
glu thr ser pro leu arg asp ala val val ile glu gln ser ser	
GAG ACC AGC CCG CTG AGG GAT GCC GTC GTC ATC GAG CAG AGC TCC	
800	
ser leu asn glu ala ser ser gly leu phe asp val phe leu arg	810
TCC CTG AAT GAG GCC AGC AGT GGC CTC TTC GAC GTC TTC CTA CGC	
820	
phe met cys his his ala val arg ile arg gly lys ser tyr val	
TTC ATG TGC CAC CAC GCC GTG CGC ATC AGG GGC AAG TCC TAC GTC	
830	
gln cys gln gly ile pro gln gly ser ile leu ser thr leu leu	840
CAG TGC CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC	
850	
cys ser leu cys tyr gly asp met glu asn lys leu phe ala gly	
TGC AGC CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG	
860	
ile arg arg asp gly leu leu leu arg leu val asp asp phe leu	870
ATT CGG CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG	

FIG. 74
(CONTINUED)

+

102/103

880

leu val thr pro his leu thr his ala lys thr phe leu arg thr
 TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC

890

leu val arg gly val pro glu tyr gly cys val val asn leu arg
 CTG GTC CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG

910

lys thr val val asn phe pro val glu asp glu ala leu gly gly
 AAG ACA GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC

920

thr ala phe val gln met pro ala his gly leu phe pro trp cys
 ACG GCT TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC

930

gly leu leu leu asp thr arg thr leu glu val gln ser asp tyr
 GGC CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC

940

ser ser tyr ala arg thr ser ile arg ala ser val thr phe asn
 TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT GTC ACC TTC AAC

950

arg gly phe lys ala gly arg asn met arg arg lys leu phe gly
 CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG

960

val leu arg leu lys cys his ser leu phe leu asp leu gln val
 GTC TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG

970

asn ser leu gln thr val cys thr asn ile tyr lys ile leu leu
 AAC AGC CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG

980

leu gln ala tyr arg phe his ala cys val leu gln leu pro phe
 CTG CAG GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT

990

his gln gln val trp lys asn pro thr phe phe leu arg val ile
 CAT CAG CAA GTT TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC

1000

ser asp thr ala ser leu cys tyr ser ile leu lys ala lys asn
 TCT GAC ACG GCC TCC CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC

1010

ala gly met ser leu gly ala lys gly ala ala gly pro leu pro
 GCA GGG ATG TCG CTG GGG GCC AAG GGC GCC GCC GGC CCT CTG CCC

1020

ser glu ala val gln trp leu cys his gln ala phe leu leu lys
 TCC GAG GCC GTG CAG TGG CTG TGC CAC CAA GCA TTC CTG CTC AAG

1030

leu thr arg his arg val thr tyr val pro leu leu gly ser leu
 CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC

1040

1050

1060

1070

1080

1090

FIG. 74
(CONTINUED)

+

```

1100
arg thr ala gln thr gln leu ser arg lys leu pro gly thr thr
AGG ACA GCC CAG ACG CAG CTG AGT CGG AAG CTC CCG GGG ACG ACG

1120
leu thr ala leu glu ala ala ala asn pro ala leu pro ser asp
CTG ACT GCC CTG GAG GCC GCA AAC CCG GCA CTG CCC TCA GAC

1130      1132
phe lys thr ile leu asp OP
TTC AAG ACC ATC CTG GAC TGA TGGCCACCCGCCACAGCCAGGCCAGAGCAGA

CACCAAGCAGCCCTGTCACGCCGGCTCTACGTCCCAGGGAGGGAGGGGCCACACCC
AGGCCCGCACCGCTGGAGTCTGAGGCCTGAGTGAGTGTTGGCCAGGCCTGCATGTCC
GGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTC
CAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTGGCTCCACCCCCAGGCCAG
CTTTTCYTCACCAGGAGCCCGGTTCCACTCCCCACATAGGAATAGTCCATCCCCAGATT
CGCCATTGTTACCCYTCGCCCTGCCYTCCTTGCCCTTCCACCCCCACCATCCAGGTGGA
GACCCCTGAGAAGGACCCCTGGGAGCTCTGGAAATTGGAGTGACCAAAGGTGTGCCCTGTA
CACAGGGCAGGACCCCTGCACCTGGATGGGGTCCCTGTGGGTCAAATTGGGGGAGGTGC
TGTGGGAGTAAAATACTGAATATATGAGTTTCAGTTTGRAAAAAAAAAAAAAAAA
AAAAAAAAAA

```

FIG. 74
(CONTINUED)

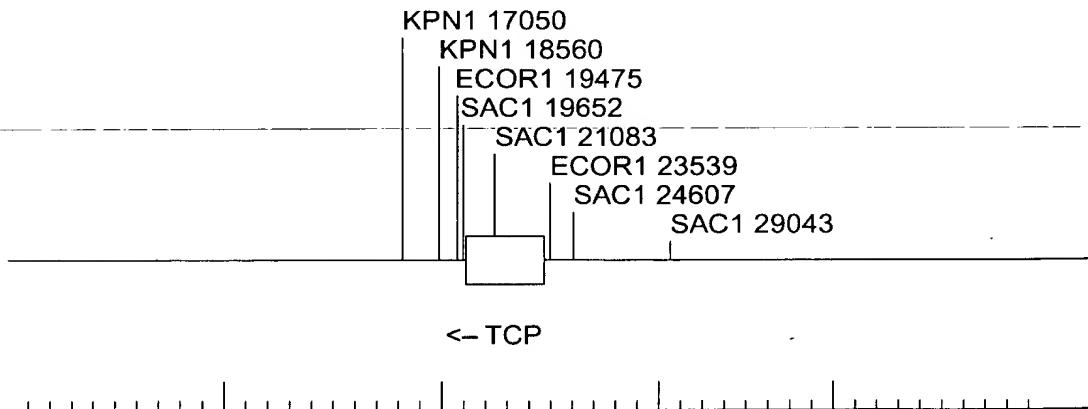


FIG. 75